Acetate metabolism in Geobacillus thermoglucosidasius and strain engineering for enhanced bioethanol production

Hills, Christopher

Award date:
2015

Awarding institution:
University of Bath

Acetate metabolism in *Geobacillus thermoglucosidasius* and strain engineering for enhanced bioethanol production

Christopher A. Hills

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Biology and Biochemistry

September 2014

Acknowledgements

First and foremost I would like to thank my supervisor Professor Michael Danson for the amazing support and guidance throughout my academic career. Without his guidance I would not be the scientist I am today. I would also like to thank Professor David Leak and Dr David Hough for their input and supervision during the project. I am also grateful for the financial contributions from the University of Bath, the BBSRC, and TMO Renewables Ltd.

I would like to thank everyone who has helped and contributed to my project. Particularly, Dr Jonathan Extance and Dr Nia Marrott for essentially teaching me everything I know! I would also like to thank everybody else in the Danson and Leak labs, including: Giannina Espina, Carolyn Williamson, Chris Vennard, Dr Dalal Binjawhar, Dr Steve Bowden, Dr Jeremy Bartosiak-Jentys, Lisa Buddrus, Dr Matthew Styles, Dr Christopher Ibenegbu, Dr Shyam Masakapalli, Beata Lisowska, Micaela Chacon, Emanuele Kendrick, Dr Leann Bacon, Ali Hussein, Luke Williams, Alexandria Holland and Dr Charlotte Bennett. You were all invaluable not only because of the wealth of experience you shared but because you made the lab such a fun and enjoyable place to be. I would also like to thank my project students for their contribution to the work: Kirsty Dundas, Kirstie Cleary, Hristina Ivanova, Mary Lunson, Eilidh Livingstone and Nick Wright.

I am also grateful for the help received from the team members at TMO Renewables Ltd who provided time and resources to run experiments. From TMO I would like to thank Dr Kirstin Eley, Dr Roger Cripps, Dr Steve Martin, Dr Gareth Cooper, Dr Alex Pudney, Dr Joanne Neary, Dr Matthew Lord and all other members of TMO Renewables Ltd for their time, support and contribution to this project.

I would like to thank my parents for supporting me during my time at the University of Bath. Lastly, special recognition must go to Lisa for all the encouragement and support she has provided during my PhD.

“Nature's laws govern which things can be done, and which can't. The trouble is, when we set out to do something, we don't always know which of these categories it's in.”

Donald Simanek (1936-)

Table of Contents

Abstract ix

List of abbreviations xi

Chapter 1 General introduction 1

1.1 Biofuel production 1

1.2 Thermophilic micro-organisms 3

1.3 *Geobacillus thermoglucosidasius* and TMO Renewables Ltd 5

1.4 Sugar metabolism 7

1.4.1 Glucose metabolism 7

1.4.2 Xylose metabolism 8

1.5 Acetate metabolism 10

1.5.1 Acetate dissimilation 10

1.5.2 Acetate assimilation 10

1.5.3 The acetate 'switch' 11

1.6 Carbon catabolite control 12

1.7 Reducing acetate production 15

1.7.1 Directly disrupting acetate production pathways 15

1.7.2 Alternative acetate generating pathways 16

1.7.3 Alternative approaches to reduce acetate production 16

1.8 Aims and objectives 17

1.9 Project outline 18

Chapter 2 Materials and Methods 20

2.1 Materials 20

2.2 Bacterial strains and plasmids 20

2.3 Bacterial growth media 21

2.3.1 Lysogeny Broth (LB) 21

2.3.2 Yeast Extract Nutrient Broth (YENB) 21

2.3.3 Super Optimal broth with Catabolite repression (SOC) 21
2.3.4 Soy Peptone Yeast Extract (No Glycerol) (2SPYNG) 21
2.3.5 Tryptone Soya Broth (TS) 21
2.3.6 Modified Two Tryptone Yeast Extract (TTY) 22
2.3.7 Glycerol Seed Culture 22
2.3.8 Modified Urea Salts Medium (USM) 22
2.3.9 Trace Elements 22
2.4 Glycerol stocks 22
2.5 *G. thermoglucosidasius* growth conditions 23
2.5.1 Tube fermentation of *G. thermoglucosidasius* 23
2.5.2 Bench-top fermenter growth of *G. thermoglucosidasius* 23
2.5.3 Analysis of fermentation products 24
2.6 Molecular biology 25
2.6.1 *G. thermoglucosidasius* genomic DNA extraction 25
2.6.2 Polymerase chain reaction 25
2.6.3 Agarose gel electrophoresis 26
2.6.4 pGEM®-T easy vector cloning 26
2.6.5 Ethanol precipitation 26
2.6.6 Preparation of electrocompetent *E. coli* 27
2.6.7 Transformation of electrocompetent *E. coli* 27
2.6.8 Transformation of chemically competent *E. coli* 27
2.6.9 Preparation of electrocompetent *G. thermoglucosidasius* 27
2.6.10 Transformation of electrocompetent *G. thermoglucosidasius* 28
2.6.11 Blue/white colony screening 28
2.6.12 Colony PCR screen 28
2.6.13 Purification of plasmid DNA 29
2.6.14 Restriction digest 29
2.6.15 Sequencing 29

2.6.16 Creation and isolation of gene knockouts in *G. thermoglucosidasius* 30
2.7 Protein biochemistry 30
2.7.1 Protein expression 30
2.7.2 *E. coli* cell extract preparation 30
2.7.3 *G. thermoglucosidasius* cell extract preparation 31
2.7.4 Nickel-affinity chromatography 31
2.7.5 Protein estimation 31
2.7.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis 32
2.7.7 Measurement of enzyme activities 32
Chapter 3 Enzyme activities in native cell lysates 36
3.1 Introduction 36
3.2 Methods 38
3.2.1 *G. thermoglucosidasius* growth 38
3.2.2 Enzyme assays 39
3.3 Results 40
3.3.1 Phosphotransacetylase 40
3.3.2 Acetate kinase 44
3.3.3 Comparison of enzyme activities throughout a fermentation run 46
3.4 Discussion 49
Chapter 4 Characterisation of recombinant phosphotransacetylase 52
4.1 Introduction 52
4.2 Methods 53
4.2.1 Cloning 53
4.2.2 Protein purification 55
4.2.3 Enzyme characterisation 56
4.2.4 Mass spectrometric analysis 57
4.3 Results 59
4.3.1 Cloning 59

4.3.2 Recombinant expression and purification 59
4.3.3 Mass spectrometry 61
4.3.4 Cloning of full-length and truncated phosphotransacetylase species 65
4.3.5 Characterisation of the two forms of recombinant phosphotransacetylase 67
4.3.6 Which protein species is native to *G. thermoglucosidasius*? 75
4.3.7 Substrate specificity 78
4.4 Discussion 78
Chapter 5 Characterisation of recombinant acetate kinase 81
5.1 Introduction 81
5.2 Methods 82
5.2.1 Cloning 82
5.2.2 Protein purification 83
5.2.3 Enzyme characterisation 83
5.3 Results 83
5.3.1 Cloning 83
5.3.2 Recombinant expression and purification 84
5.3.3 Characterisation of the recombinant acetate kinase 85
5.4 Discussion 87
Chapter 6 Adaptation of the pTMO31 integration vector and deletion of *pta* 89
6.1 Introduction 89
6.2 Methods 91
6.2.1 Effects of kanamycin concentration on pTMO31 transformation 91
6.2.2 Creation of pUB31 91
6.2.3 Transformation efficiencies 92
6.2.4 Creation of a *pta* knockout construct 92
6.2.5 Isolation of mutant strains with *pta* deletion 93
6.2.6 Enzyme assays 93
6.2.7 Tube fermentations 93

6.2.8 Bench-top fermentations 93
6.3 Results 94
6.3.1 Adaptation of the integration vector 94
6.3.2 Deletion of *pta* and strain characterisation 95
6.4 Discussion 105
Chapter 7 Could a promiscuous phosphotransacylase compensate for the loss of phosphotransacetylase? 108
7.1 Introduction 108
7.2 Methods 110
7.2.1 Cloning 110
7.2.2 Protein purification 112
7.2.3 Enzyme characterisation 112
7.2.4 Creation of *ptp* and *ptb* knockout constructs 113
7.2.5 Isolation of mutant strains with gene deletion 115
7.2.6 Enzyme assays 115
7.2.7 Tube fermentations 115
7.2.8 Bench-top fermentations 115
7.2.9 Voges-Proskauer test 116
7.3 Results 116
7.3.1 Cloning 116
7.3.2 Expression and purification of recombinant enzymes 118
7.3.3 Characterisation of the recombinant phosphotransacylases 120
7.3.4 Substrate specificity of the recombinant butyrate kinase 127
7.3.5 Deletion of *ptp* and *ptb*, and strain characterisation 128
7.4 Discussion 140
Chapter 8 Acetyl-CoA synthetase and regulation by lysine acetylation 143
8.1 Introduction 143
8.2 Methods 146

8.2.1 Cloning 146
8.2.2 Protein purification 148
8.2.3 Enzyme characterisation 148
8.2.4 Acetyl-CoA synthetase regulation 150
8.2.5 The acetylation of phosphotransacetylase 151
8.3 Results 151
8.3.1 Cloning and recombinant expression 151
8.3.2 Characterisation of acetyl-CoA synthetases 153
8.3.3 Regulation of acetyl-CoA synthetase 156
8.3.4 The regulation of phosphotransacetylase by lysine acetylation 157
8.4 Discussion 160
Chapter 9 General conclusion 162
Chapter 10 References 167

Abstract

Social, economic and political pressures have driven the development of renewable alternatives to fossil fuels. Biofuels, such as bioethanol, have proved to be successful alternatives. Mature technologies are crop-based, but this has brought criticism due to the conflicting use of land for fuel versus food production. Therefore, bioethanol production technologies have shifted to utilising the sugars that derive from the degradation of lignocellulosic biomass. The thermophilic, Gram-positive bacterium, *Geobacillus thermoglucosidasius*, can naturally utilise a large fraction of these sugars, and metabolic engineering has been used to create a strain that produces ethanol as the major product of fermentation. This strain, *G. thermoglucosidasius* TM242 (Δ*ldh*, Δ*pfl*, *pdh*^{up}), does however, produce small but significant quantities of acetate, an undesirable by-product of fermentation. Therefore, acetate metabolism in the *G. thermoglucosidasius* TM242 strain was the focus of this study.

During fermentation, ethanol is generated from the central metabolite acetyl-CoA through the activities of a bifunctional enzyme: aldehyde dehydrogenase/alcohol dehydrogenase (ADHE). On the other hand, acetate is generated from acetyl-CoA through catalysis by phosphotransacetylase (PTA) and acetate kinase (AK). Acetate metabolism in *G. thermoglucosidasius* TM242 was studied in this project by investigating the enzyme activities governing flux from acetyl-CoA, and the feasibility of reduced acetate production was investigated by a *pta*-deletion strategy.

This thesis reports the characterisation of PTA and AK, by studying activities from both native cell lysates and recombinantly expressed proteins. The results indicate that the activities of PTA and AK are greater than those of ADHE, suggesting that the potential metabolic flux is greater towards acetate production than to ethanol. However, the ethanol yield from *G. thermoglucosidasius* TM242 fermentations is greater than that of acetate, suggesting the existence of a regulatory mechanism controlling acetyl-CoA flux. Several possible regulatory mechanisms were studied in this project and are reported here.

The viability of creating a strain that reduces acetate accumulation, and potentially increases ethanol yields, was investigated and reported in this thesis. The gene encoding PTA was deleted from *G. thermoglucosidasius* TM242, and the resulting strain was

characterised. The Δ*pta* strain had approximately 5% of the PTA activity measured in TM242, but acetate was still generated from pentose and hexose fermentations. Additional phosphotransacylase (PTAC) enzymes were discovered in *G. thermoglucosidasius* TM242 that could catalyse the conversion of acetyl-CoA and orthophosphate to acetyl-phosphate and CoA. A series of PTAC null strains were created and analysed, the results of which indicated that phosphotransbutyrylase (PTB) could be involved in acetate production *in vivo*. It was discovered that the cell lysates of *G. thermoglucosidasius* strains carrying deletions to both *pta* and *ptb* could no longer catalyse the conversion of acetyl-CoA and orthophosphate to acetyl-phosphate and CoA. However, these strains still accumulated acetate, suggesting the presence of alternative acetate-producing pathways in this organism. In addition, *G. thermoglucosidasius* strains carrying deletions to both *pta* and *ptb* could ferment glucose but not xylose, suggesting that the production of ATP by the PTA-AK pathway is crucial for micro-aerobic growth on pentose sugars.

List of abbreviations

Abbreviation	Definition
2-CST	Two-component signal transduction system
2SPY(NG)	Soy peptone yeast extract (no glycerol)
ACS	Acetyl-CoA synthetase
AcuA	Lysine acetyltransferase
AcuC	Lysine deacetylase
ADH	Alcohol dehydrogenase
ADHE	Bifunctional AldDH/ADH enzyme
AK	Acetate kinase
AldDH	Aldehyde dehydrogenase
BK	Butyrate kinase
CAC	Citric acid cycle
CCA	Carbon catabolite activation
CcpA	Carbon catabolite protein A
CCR	Carbon catabolite repression
cre	Catabolite-response element
Crh	Catabolite repression HPr
DLS	Dynamic light scattering
EC	Enzyme commission (number)
ED	Entner-Douderoff
EMP	Embden-Meyerhof-Parnas
His-tag	Histidine tagged (protein)
HK	Histidine kinase
HPr	Phosphocarrier protein
HPrK	HPr kinase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KAT	Lysine acetyltransferase
KDAC	Lysine deacetylase
KNTase	Kanamycin nucleotidyltransferase
LB	Lysogeny broth
LDH	Lactate dehydrogenase
OD	Optical density
PDH	Pyruvate dehydrogenase
PFL	Pyruvate formate lyase
PK	Pyruvate kinase
PoxB	Pyruvate oxidase
PTA	Phosphotransacetylase
PTAC	Phosphotransacylase
PTB	Phosphotransbutyrylase
PTP	Phosphotranspropionylase
PTS	Phosphoenolpyruvate: sugar phosphotransferase
SOC	Super optimal broth with catabolite repression
TS	Tryptone soya broth
TTY	Two tryptone yeast extract
U	µmol/min
USM	Urea salts medium
YENB	Yeast extract nutrient broth
ε	Extinction coefficient

Chapter 1
General introduction

1.1 Biofuel production

Growing concerns over environmental issues, finite fossil fuel supplies and energy security have stimulated an increased interest in microbial fuel production using renewable raw materials (Kerr 2007; Stephanopoulos 2007). Biofuels such as biodiesel, biobutanol, bioethanol and biohydrogen have great potential as renewable alternatives to fossil fuels because they are derived from plant biomass: an abundant, renewable source of carbohydrates for microbial conversion to chemicals and biofuels (Alonso *et al.* 2010; Sun and Cheng 2002). Biofuel production converts the energy captured from the sun by plants through photosynthesis into chemical energy in the bonds of biologically produced molecules (Rude and Schirmer 2009). This, in contrast to fossil fuels, is a renewable process, and has the potential to be carbon neutral because the feedstocks are synthesised by recent plant-based carbon dioxide fixation.

Ethanol is the most commonly produced biofuel, with global bioethanol production surpassing 23 billion US gallons in 2013 (Renewable Fuels Association, RFA 2014). This is largely due to the high number of microbes that can naturally produce ethanol, through fermentation of cornflour (in the United States) or cane sugar (in Brazil). Mature technologies for ethanol production are thus crop-based; typical crops include sugar cane, corn, beets, wheat, sorghum, sunflower, soybean, cassava, etc. These feedstocks are rich in starch, which when processed result in the release of simple sugars that can be fermented to ethanol by microorganisms such as *Saccharomyces cerevisiae* (Sanchez and Cardona 2008).

To cover future energy demands, high-volume production of biofuels would be required. However, it has been argued that high-volume first-generation bioethanol production from food crops would be unsustainable because of the link created between food and fuel prices (Haber 2007; Tenenbaum 2008). Instead, bioethanol could be produced from lignocellulose, a naturally occurring and potentially cheap polymer found in agricultural residue (wheat straw, corn stalks, soybean residues, sugar cane bagasse), industrial waste (pulp and paper industry), forestry residues, municipal solid waste, etc. It has been

estimated that this non-food energy feedstock accounts for about 50% of the biomass on Earth (10-50 billion tons according to Claassen *et al.* (1999)). However, to release the sugars from lignocellulose, thermal, chemical and biochemical processing is required prior to fermentation by micro-organisms (Peralta-Yahya *et al.* 2012).

Lignocellulosic bioethanol costs are high due to the energy and time required to disrupt the complex matrix of interlinked polymers that form the basic structure of the plant cell wall (Alvira *et al.* 2010; Mosier *et al.* 2005; Zaldivar *et al.* 2001). The composition of lignocellulose is broadly made up of cellulose strands (a polymer of glucose molecules) cross-linked by hemi-cellulose (polymers of xylose, mannose, galactose, and arabinose), which is then further strengthened by lignin and pectin. The exact composition varies between different feedstocks, and therefore the method used to disrupt the lignocellulosic structure requires tailoring and optimising prior to mass ethanol production; however, methods typically include a physical pretreatment stage followed by an extensive enzymatic hydrolysis stage (Jordan *et al.* 2012; Lynd *et al.* 2008). The pretreatment process softens and prepares the biomass for enzymatic hydrolysis; common methods include mechanical reduction in biomass particle size by milling, steam explosion, hydrothermolysis and acid or base hydrolysis (Mosier *et al.* 2005).

In addition, second-generation bioethanol production costs are high because the degradation of lignocellulosic biomass generates a diverse set of sugars that many conventional yeast strains cannot fully ferment (Mosier *et al.* 2005). Conventional yeast strains can typically ferment C6 monomers (glucose and mannose); however, most cannot ferment C5 monomers (xylose and arabinose), or short-chain polysaccharides (cellobiose and xylobiose) that also result from the degradation of lignocellulose. Full use of these carbohydrates is required if a lignocellulosic bioethanol process is to be economically viable.

An ideal microbe for lignocellulosic bioethanol production would have many desirable traits including balanced growth on both hexose and pentose sugars, tolerance to pre-treatment inhibitors, the ability to produce stable hydrolytic enzymes, maximum ethanol yield and production rates, solvent tolerance, and the ability to withstand process fluctuations (Elkins *et al.* 2010). The potential of designing and constructing a microbe possessing these traits has become a reality with the recent progress in metabolic engineering, and the study of biological systems as a whole (Peralta-Yahya *et al.* 2012; Stephanopoulos 2007).

Consequently, various groups have engineered metabolic pathways in an attempt to extend the substrate utilisation of particular microbes (Elkins *et al.* 2010; Jordan *et al.* 2012; Sanchez and Cardona 2008). Maurizio Bettiga *et al*. (2009) describe the development of a *S. cerevisiae* strain expressing a fungal pathway for the utilisation of L-arabinose and D-xylose, while Brat *et al*. (2009) introduced a xylose isomerase from *Clostridium phytofermentans* into *S. cerevisiae*, enabling the yeast cells to metabolise D-xylose. An alternative approach is to use microbes that naturally have the ability to ferment a wide range of substrates, and engineer in the metabolic pathways required to produce ethanol. For example, thermophilic organisms, such as some species of *Clostridium*, *Thermoanaerobacter* and *Geobacillus*, often can ferment a wide range of carbohydrates in addition to possessing other traits suitable for industrial bioethanol production (Barnard *et al.* 2010; Bhalla *et al.* 2013; Kumar *et al.* 2009; Lin and Xu 2013; Taylor *et al.* 2009).

1.2 Thermophilic micro-organisms

Thermophiles belong to a sub-category of extremophilic microorganisms that grow at temperatures between 40°C and 70°C, and they have many properties that are advantageous in an industrial bioethanol process. Generally, thermophiles are robust organisms producing stable enzymes, and are often able to withstand changes in environmental conditions, such as pH and temperature (Barnard *et al.* 2010; Hild *et al.* 2003). Thermophiles are often able to ferment the pentose and hexose sugar fraction of biomass; they are also able to ferment hydrolysate materials, and in some cases polymeric precursors or structurally-complex polycarbohydrates such as cellulose (Sommer *et al.* 2004; Zaldivar *et al.* 2001). This ability to utilise a wide substrate range is particularly important for the production of second-generation biofuels (Taylor *et al.* 2009). Furthermore, industrial fermentations at high temperatures (>50°C) are less prone to microbial contamination (Barnard *et al.* 2010; Lin and Xu 2013; Taylor *et al.* 2009). Therefore, the addition of antibiotics, which can add significant costs and have negative environmental consequences, is not required. The operation of a bioprocess at higher temperatures is also beneficial because high temperatures accelerate the chemical reaction rate, reduce the required energy input, promote better solubility and efficient mixtures of substrate, and facilitate the removal of volatile products (Barnard *et al.* 2010; Lin and Xu 2013).

Thermophiles, therefore, not only provide a valuable source of thermostable enzymes for biotechnology, such as hydrolases (e.g. cellulases, xylanases, and proteases) and DNA

polymerases (e.g. *Taq*), but they are also potentially useful microbial cellular factories. To date, thermophilic ethanol production has been reported using *Clostridium thermocellum* (Argyros *et al.* 2011), *Thermoanaerobacterium saccharolyticum* (Shaw *et al.* 2008) and *Geobacillus thermoglucosidasius* (Cripps *et al.* 2009). In addition, n-butanol and isobutanol have been produced using *T. saccharolyticum* and *G. thermoglucosidasius*, respectively (Bhandiwad *et al.* 2014; Lin *et al.* 2014). Furthermore, the hyperthermophile *Pyrococcus furiosus* has been reported to produce 3-hydroxypropionic acid (Keller *et al.* 2013).

Thermophilic Clostridia are potentially useful candidates for use in a biofuel production process through their ability to degrade and ferment crystalline cellulose to a mixture of solvents (e.g. ethanol, butanol and isopropanol). In some *Clostridium* species, this ability to degrade cellulose is via the expression of a diverse set of hydrolase enzymes that form a multi-enzyme complex called a cellulosome (Barnard *et al.* 2010; Lin and Xu 2013; Taylor *et al.* 2009). Another group of thermophilic anaerobes belonging to the genus *Thermoanaerobacter* can ferment both glucose and xylose to ethanol (Lacis and Lawford 1991). Similarly, bacteria belonging to the *Geobacillus* genus can also produce ethanol but can ferment a wider range of substrates than *Thermoanaerobacter*; for example, some *Geobacillus* species are able to ferment glucose, xylose, and arabinose between temperatures of 55°C and 70°C (San Martin *et al.* 1992), while other *Geobacilli* have been identified that can degrade more complex polycarbohydrates such as xylan (Wu *et al.* 2006).

Despite the advantages associated with using thermophiles for biofuel production, there are some limitations that currently prevent an efficient, economically viable process. High ethanol yields are typically lacking because thermophilic fermentation often results in a mixture of products, such as additional solvents and organic acids (Chang and Yao 2011; Dien *et al.* 2003). The formation of organic acid by-products is not only an inefficient use of the carbon source, but acid production can also lead to a retardation of cellular growth. Tolerance to ethanol is also typically low in thermophilic ethanologens compared to mesophilic ethanologens such as *Saccharomyces cerevisiae* or *Zymomonas mobilis* (Ben-Bassat *et al.* 1981; Demain *et al.* 2005; Zeikus *et al.* 1981). Genetic engineering can be used to overcome these limitations; for example, a *Thermoanaerobacter ethanolicus* strain has been developed to tolerate up to 8% (v/v) ethanol (Burdette *et al.* 2002), while the

Geobacillus thermoglucosidasius strain M10EXG has also been reported to tolerate ethanol up to 10% (v/v) (Fong *et al.* 2006).

Until recently, progress in using thermophiles has been slow, largely due to the barriers hampering the genetic manipulation of these organisms. These barriers include the difficulty of bacterial transformation, either due to the lack of reported techniques or due to the hindrance from the physiological features of the cell. Thermophilic bacteria tend to have a tough cell envelope, many form endospores, and often they have a plasma membrane that is only weakly permeable (Averhoff 2004; Culha *et al.* 2008; Silhavy *et al.* 2010). The lack of well-established genetic toolkits has limited the rate at which thermophiles have been employed in industrial processes. Suitable plasmids must have a thermostable origin of replication, and the use of thermostable and efficient marker-genes are required for genetic manipulation to be successful.

To overcome these barriers, several recent techniques and tools have been developed. A host of different transformation techniques have been successfully reported for thermophiles, including electroporation, conjugation and sonoporation (Cesar *et al.* 2011; Lin *et al.* 2010; Olson and Lynd 2012b). Advances have also been made in the genetic toolkit available for certain thermophiles. A number of thermostable plasmids have been reported that allow the expression of both foreign and native genes in thermophilic hosts (Cava *et al.* 2009; Taylor *et al.* 2008; 2009). In addition, plasmids have been developed that allow chromosomal deletion and/or insertion of genes in certain thermophiles (Cripps *et al.* 2009; Olson and Lynd 2012a; Tripathi *et al.* 2010). Thermostable antibiotic selection markers and counter-selection systems have also aided the manipulation of thermophiles (Argyros *et al.* 2011; Cava *et al.* 2009; Liao and Kanikula 1990; Suzuki *et al.* 2012; Tripathi *et al.* 2010). The rapid development of the techniques and genetic toolkits available to manipulate thermophiles has led to the establishment of many new biotechnology companies exploiting thermophiles to create chemicals and biofuels.

1.3 *Geobacillus thermoglucosidasius* and TMO Renewables Ltd

Geobacillus thermoglucosidasius is a facultatively anaerobic, rod-shaped, Gram-positive, endospore-forming bacterium (Nazina *et al.* 2001). *G. thermoglucosidasius* is capable of growth at temperatures between 40°C and 70°C and can ferment hexose and pentose sugars to generate lactate, formate, acetate and ethanol as products. Under oxygen-limited

growth, the major product from the *G. thermoglucosidasius* NCIMB 11955 wild-type strain is lactate, as a method of regenerating NAD^+ consumed during glycolysis (Figure 1.1). The potential for using *G. thermoglucosidasius* as a platform for lignocellulosic bioethanol production has been explored by the biotechnology company TMO Renewables Ltd (TMO), Surrey, UK (relaunched in 2014 as ReBio Technologies Ltd).

A suitable genetic toolkit and transformation procedures have been established for *G. thermoglucosidasius*, creating the ability to over-express genes in this host using the plasmid pUCG18, and to manipulate the genomic DNA by deletion or insertion using the plasmid pTMO31 (Taylor *et al.* 2008; Cripps *et al.* 2009). It has been reported that an engineered strain with deletions to genes encoding lactate dehydrogenase (LDH) and pyruvate formate lyase (PFL), and the up-regulated expression of the pyruvate dehydrogenase (PDH) complex (Figure 1.1), could generate ethanol as the major product from both hexose and pentose sugars (Cripps *et al.* 2009). The resulting *G. thermoglucosidasius* strain (Δ*ldh*, Δ*pfl*, pdh^{up}), known as TM242, has been demonstrated to be capable of producing ethanol at yields in excess of 90% of the theoretical maximum (0.41-0.44 g of ethanol / g of glucose consumed), and at the start of this project was the current production strain utilised by TMO.

The *G. thermoglucosidasius* TM242 strain has increased carbon flow towards acetyl-CoA production, and despite the additional consumption of NAD^+ through the activity of PDH, the reduction of acetyl-CoA to ethanol is sufficient to balance the redox state of the cell (Figure 1.1). This is achieved by the activity of the bifunctional enzyme aldehyde dehydrogenase/alcohol dehydrogenase (ADHE), which catalyses the conversion of acetyl-CoA to ethanol with the concomitant re-generation of NAD^+, previously consumed during glycolysis (Extance *et al.* 2013). Acetate is the only other significant by-product of *G. thermoglucosidasius* TM242 fermentation, and is 3.6-fold greater at peak ethanol concentration when fermenting xylose compared to glucose (Cripps *et al.* 2009). The increase in acetate production from xylose fermentation suggests that the differences between glycolysis and the pentose phosphate pathway of sugar metabolism are significant factors in influencing downstream carbon flux. The production of this unwanted organic acid by-product from *G. thermoglucosidasius* TM242 fermentation is the focus of the study reported in this thesis.

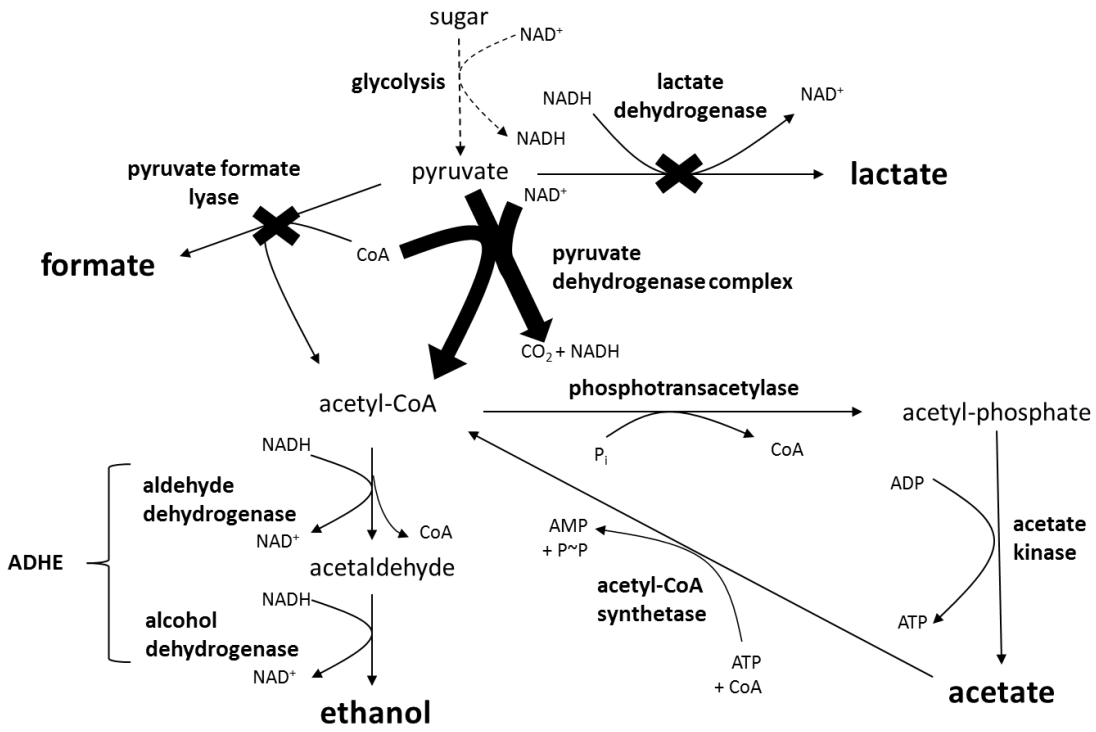


Figure 1.1 Schematic representation of the major metabolic pathways surrounding ethanol production in *Geobacillus thermoglucosidasius*. The crosses indicate gene deletions, and bold arrows represent up-regulation of gene expression in the *G. thermoglucosidasius* TM242 strain. Key to abbreviations: ADHE, aldehyde dehydrogenase/alcohol dehydrogenase; P_i, orthophosphate; P~P, pyrophosphate.

1.4 Sugar metabolism

1.4.1 Glucose metabolism

There are two major pathways for the catabolism of glucose to pyruvate: the Embden-Meyerhof-Parnas (EMP) pathway of glycolysis (Cooper 1978) and the Entner-Doudoroff (ED) pathway (Entner and Doudoroff 1952). Considering that only the genes involved with EMP glycolysis have been discovered in *G. thermoglucosidasius* TM242, the ED pathway will not be considered here (TMO, unpublished data). In classical EMP metabolism, glucose is converted to fructose-1,6-bisphosphate, which is cleaved to glyceraldehyde-3-phosphate and dihydroxyacetone phosphate. A second molecule of glyceraldehyde-3-phosphate is generated from dihydroxyacetone phosphate, before being converted to pyruvate in a five-step reaction sequence (Figure 1.2). For every molecule of glucose metabolised, two molecules of pyruvate, ATP, and NADH are generated.

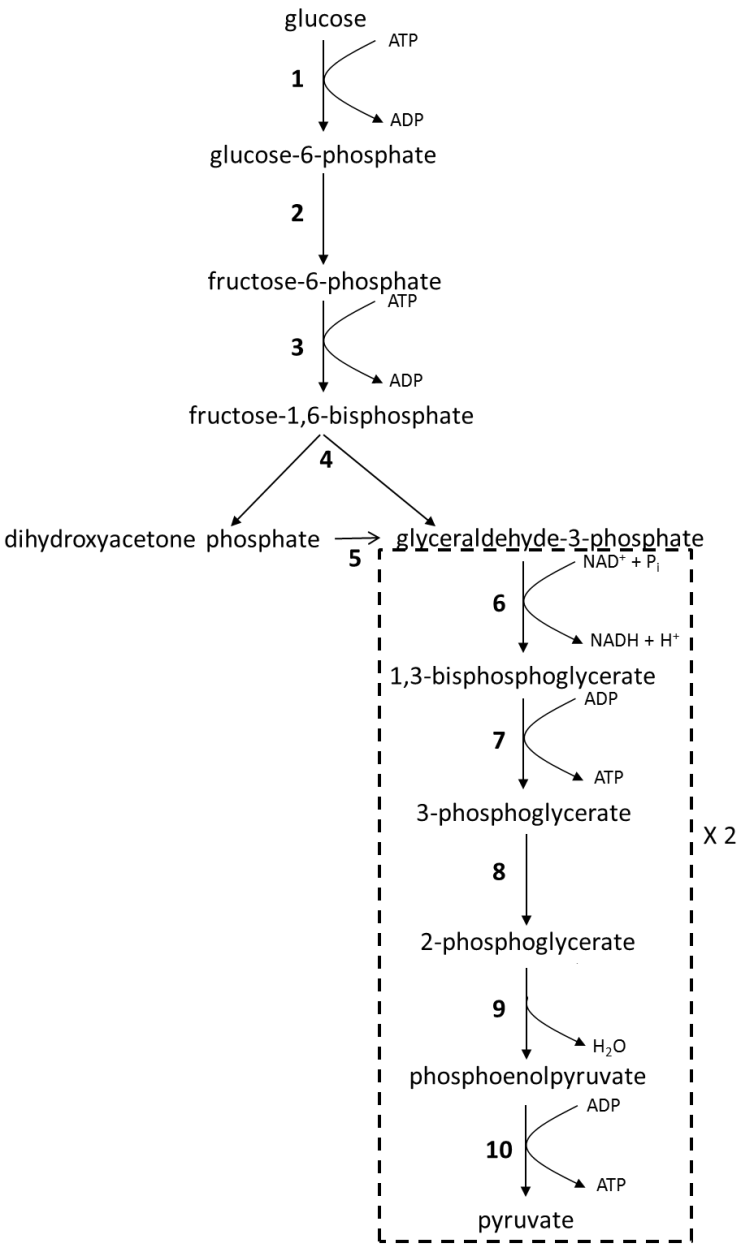


Figure 1.2 Schematic representation of the Embden-Meyerhof-Parnas pathway. The net reaction is:

$$\boldsymbol{glucose + 2\ ADP + 2\ NAD^{+} + 2\ P_i \rightarrow 2\ pyruvate + 2\ ATP + 2\ NADH + 2\ H^{+}}$$

Enzymes are denoted by numbers: 1, hexokinase; 2, phosphoglucose isomerase; 3, phosphofructokinase; 4, fructose bisphosphate aldolase; 5, triosephosphate isomerase; 6, glyceraldehyde-3-phosphate dehydrogenase; 7, phosphoglycerate kinase; 8, phosphoglycerate mutase; 9, enolase; 10, pyruvate kinase. The last five reactions indicated by the dashed box occur twice for every molecule of glucose catabolised, as the dihydroxyacetone phosphate is converted to a second molecule of glyceraldehyde-3-phosphate.

1.4.2 Xylose metabolism

There are three known pathways for the catabolism of xylose to pyruvate in prokaryotes: the isomerase pathway, the Weimberg pathway (Weimberg 1961) and the Dahms pathway (Dahms 1974). It appears as though *G. thermoglucosidasius* TM242 utilises the isomerase pathway to catabolise xylose, due to the presence of a gene annotated as

encoding a xylose isomerase, and the absence of xylose dehydrogenase, xylonolactonase and xylonate dehydrogenase required for the Weimberg and Dahms pathways (TMO, unpublished data). The xylose isomerase catalyses the conversion of xylose directly to xylulose, which once phosphorylated can enter the pentose phosphate pathway which coverts xyulose-5-phosphate to pyruvate (Figure 1.3). The fructose-6-phosphate and glyceraldehyde-3-phosphate enter glycolysis at different stages, but are still converted to pyruvate as shown in Figure 1.2. For every three molecules of xylose consumed, five molecules of pyruvate, ATP, and NADH are generated. This results in a net ATP yield of 1.67 molecules of ATP generated per molecule of xylose catabolised, lower than the ATP yield from glycolysis (2 ATPs per glucose consumed). The lower ATP yield from xylose metabolism may be a significant factor in influencing acetate production, because this high-energy metabolite is generated from the conversion of acetyl-CoA to acetate by the activities of phosphotransacetylase (PTA, EC 2.3.1.8) and acetate kinase (AK, EC 2.7.2.1) (Figure 1.1).

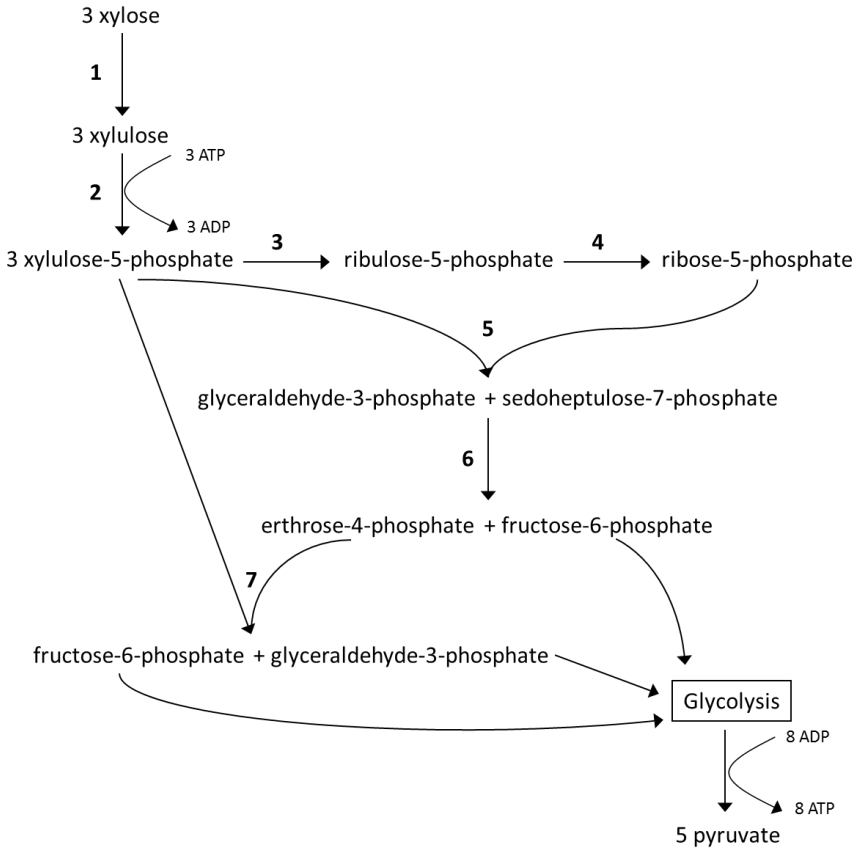


Figure 1.3 Schematic representation of xylose catabolism in *G. thermoglucosidasius*. The net reaction is:

$$\mathbf{3}\ \boldsymbol{xylose} + \mathbf{5}\ \boldsymbol{ADP} + \mathbf{5}\ \boldsymbol{P_i} + \mathbf{5}\ \boldsymbol{NAD^+} \rightarrow \mathbf{5}\ \boldsymbol{pyruvate} + \mathbf{5}\ \boldsymbol{ATP} + \mathbf{5}\ \boldsymbol{NADH} + \mathbf{5}\ \boldsymbol{H^+}$$

Enzymes are denoted by numbers: 1, xylose isomerase; 2, xylulose kinase; 3, ribulose-5-phosphate epimerase; 4, ribulose-5-phosphate isomerase; 5, transketolase; 6, transaldolase; 7, transketolase.

1.5 Acetate metabolism

1.5.1 Acetate dissimilation

Sugar metabolism to acetyl-CoA via glycolysis consumes NAD^+ and results in a depletion of free CoA. In the absence of oxygen, for example during fermentative growth, the citric acid cycle (CAC) cannot complete the oxidation of acetyl-CoA to carbon dioxide. Therefore, ATP, CoA and NAD^+ must be generated via alternative pathways. Under these conditions, acetogenesis, the excretion of acetate into the environment, occurs (Elmansi and Holms 1989; Majewski and Domach 1990). The energy requirement for the cell can be met by ATP production via substrate-level phosphorylation in glycolysis and in the PTA-AK pathway (Figure 1.1). The conversion of acetyl-CoA to acetate, via the PTA-AK pathway, generates two ATP molecules per glucose, but consumes no reducing equivalents. The reduction of acetyl-CoA to ethanol, catalysed by ADHE, sacrifices energy-production but consumes reducing equivalents. Therefore, a cell can balance its requirements to regenerate NAD^+ with its need for energy by modulating the quantities of ethanol and acetate it produces.

Acetogenesis also occurs when the influx of carbon into the cells exceeds the CAC's capacity and that of other central metabolic pathways. Therefore, acetate excretion occurs both anaerobically during mixed-acid fermentation, and aerobically when growth on an excess carbon source inhibits respiration, a behaviour called the Crabtree effect (Crabtree 1929). Acetate metabolism has been extensively studied in *E. coli* and *Bacillus subtilis*. Both organisms utilise the PTA-AK pathway to excrete acetate during exponential growth. Unlike *E. coli*, *B. subtilis* diverts some of the pyruvate to uncharged acetoin or 2,3-butanediol to prevent the over-acidification of the environment (Speck and Freese 1973). During the late exponential growth phase, acetoin undergoes oxidative dissimilation to acetaldehyde and finally acetate (Huang *et al.* 1999; Lopez *et al.* 1975).

1.5.2 Acetate assimilation

During the stationary phase, when the carbon source has become depleted, *E. coli* and *B. subtilis* can assimilate and activate acetate to acetyl-CoA via the AMP-forming acetyl-CoA synthetase (ACS, EC 6.2.1.1) (Grundy *et al.* 1993b; Wolfe 2005). In *E. coli*, at acetate concentrations above 30mM, acetyl-CoA can be generated by the reversal of the PTA-AK pathway, but there is evidence suggesting *B. subtilis* does not utilise the PTA-AK pathway for growth on acetate (Grundy *et al.* 1993b). The synthesis of acetyl-CoA by the AMP-forming

ACS is catalysed via an ordered Bi Uni Uni Bi ping-pong mechanism (Starai and Escalante-Semerena 2004a). ATP is bound to the enzyme first, then acetate, which leads to the formation of enzyme-bound acetyl-AMP and the release of pyrophosphate. Next, the CoA binds to the enzyme, which catalyses the formation of acetyl-CoA before the enzyme releases acetyl-CoA followed by AMP. The formation of the acetyl-AMP intermediate is irreversible *in vivo* due to the formation of pyrophosphate, which would subsequently be cleaved by pyrophosphatases within the cell. Therefore, the AMP-forming ACS is considered to be only an acetate-uptake system *in vivo*.

1.5.3 The acetate 'switch'

The 'switch' from acetate dissimilation to acetate assimilation in *E. coli* has been extensively reviewed by Wolfe (2005), and is defined by the point at which the rate of acetate dissimilation is equal to the rate of acetate assimilation. This 'switch' is largely dictated by controlling the initiation of *acs* transcription. During exponential growth, *acs* transcription remains low, and acetate is generated. Transcription of *acs* is induced during the transition to the stationary phase. The regulation of *acs* transcription appears to be quite complex, involving at least two promoters, two sigma factors, the transcription factor CRP, and two nucleoid-associated proteins: IHF and FIS. In addition, the activity of the AMP-forming ACS is regulated post-translationally by an acetylation-deacetylation system (Figure 1.4), in which a lysine acetyltransferase (KAT) acetylates a lysyl residue at position 609, resulting in the inactivation of ACS (Starai and Escalante-Semerena 2004b). The acetylation inhibits the adenylation of acetate, but the thioester-forming step remains unaffected. An NAD^+-dependent lysine deacetylase (KDAC), CobB, reactivates ACS generating nicotinamide and 2'-*O*-acetyl-ADP-ribose as co-products (Hu *et al.* 2010).

The 'switch' to acetate assimilation has also been studied in the Gram-positive bacterium *B. subtilis*, and like in *E. coli* it appears to be dictated by the regulation of the initiation of *acs* transcription. *B. subtilis*, however, appears to have a different system to control the expression of *acs*, based on carbon catabolite repression (Grundy *et al.* 1994; Presecan-Siedel *et al.* 1999). On consumption of rapidly metabolisable sugars such as glucose, the carbon catabolite protein A (CcpA) is activated, which binds to a catabolite-response element (*cre*) site downstream of the transcription initiation bases of *acs*, thus blocking transcription (Grundy *et al.* 1994; Zalieckas *et al.* 1998). Hence, under conditions where the

glucose concentration is high, *acs* transcription is low; conversely, transcriptional inhibition will be lifted under low glucose concentrations.

In addition, a similar post-translational acetylation-deacetylation system regulates the activity of ACS in *B. subtilis* (Figure 1.4). The acetylation of ACS is catalysed by a lysine acetyltransferase, AcuA, inactivating the enzyme. However, unlike the *E. coli* NAD^+-dependent lysine deacetylase, the *B. subtilis* AcuC was determined to be an NAD^+-independent lysine deacetylase (Gardner *et al.* 2006). The AcuA and AcuC enzymes were discovered to be expressed from a *B. subtilis* operon annotated as *acuABC* (acetoin utilisation), and the AcuB protein was homologous to proteins with unknown function.

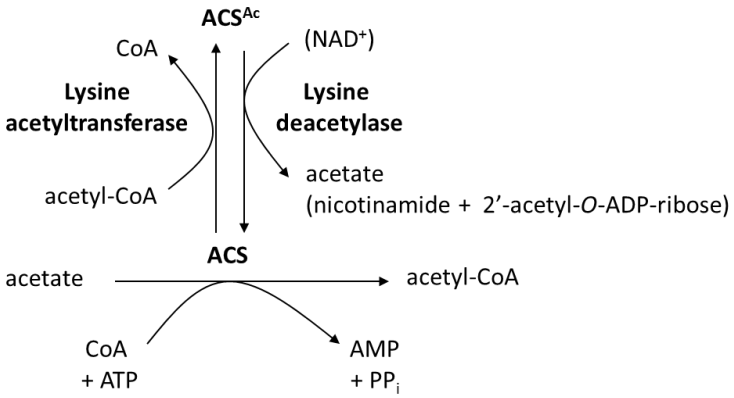


Figure 1.4 Schematic representation of acetyl-CoA synthetase regulation by an acetylation-deacetylation system. Acetyl-CoA synthetase is active in its non-acetylated form (ACS) and catalyses the conversion of acetate, CoA and ATP to acetyl-CoA, AMP and pyrophosphate (PP_i). An acetyl-CoA-consuming lysine acetyltransferase can acetylate ACS, converting it to an inactive form (ACS^{AC}). A lysine deacetylase can reactivate ACS by cleaving the acetyl-group from the lysyl residue, which releases acetate, or an NAD^+-dependent lysine deacetylase can remove the acetyl-group generating nicotinimide and 2'-*O*-acetyl-ADP-ribose. The NAD^+-dependent lysine deacetylase reaction is in parentheses.

1.6 Carbon catabolite control

Bacteria have developed sophisticated mechanisms to survive and adapt to environmental changes. Carbon catabolite repression (CCR) allows bacteria to rapidly assimilate a preferred carbon source when they are exposed to more than one carbohydrate. This regulatory system is a mechanism by which the cell can maximise its efficiency of carbon utilisation and also regulate other metabolic processes (Deutscher 2008). Many bacteria use the phosphoenolpyruvate:sugar phosphotransferase system (PTS) for the uptake of various sugars, and this is also involved in CCR in certain Gram-positive bacteria, including *B. subtilis*. The uptake of rapidly metabolisable PTS-sugars increases the concentration of

glycolytic intermediates such as fructose-1,6-bisphosphate (FBP), which stimulates the activity of the HPr kinase (HPrK). The HPrK catalyses the ATP-dependent phosphorylation of the phosphocarrier protein HPr, and its homologue Crh (catabolite repression HPr), at the regulatory Ser-46 site. These phosphorylated forms of HPr and Crh interact with CcpA, and this complex binds to specific DNA targets, the *cre* sites, to elicit the carbon catabolite response in the cell (Figure 1.5) (Deutscher 2008; Fujita 2009; Presecan-Siedel *et al.* 1999).

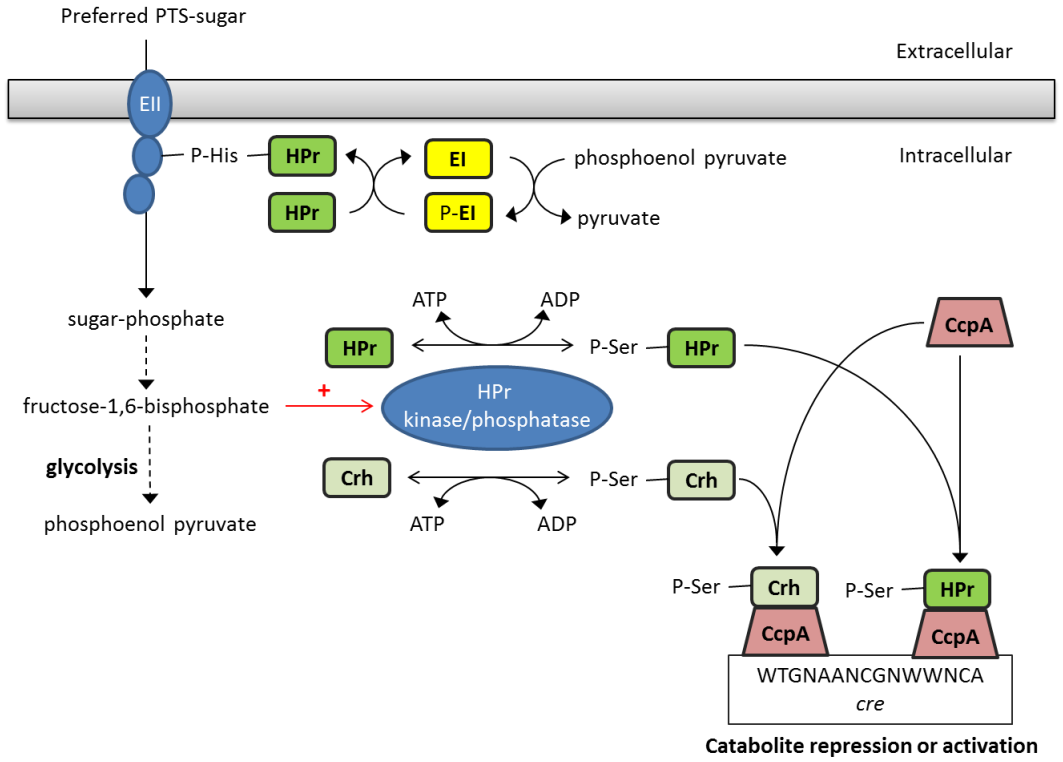


Figure 1.5 Schematic representation of the mechanism of carbon catabolite repression and activation in Gram-positive bacteria. The phosphoenol pyruvate:sugar phosphotransferase system (PTS) is a major carbohydrate transport system in bacteria. The PTS catalyses the phosphorylation of preferred sugars, such as glucose, fructose or mannose, during their translocation across the cell membrane. The mechanism involves the transfer of a phosphoryl group from phosphoenol pyruvate, via enzyme I (EI), to a conserved N-terminal His on HPr. In Gram-positive bacteria an increase in fructose-1,6-bisphosphate concentration triggers ATP-dependent HPr kinase/phosphatase-catalysed phosphorylation of HPr and Crh at Ser-46. Only the seryl-phosphorylated forms of HPr and Crh are capable of binding to CcpA. The P-Ser-HPr/CcpA and P-Ser-Crh/CcpA complexes can bind to the catabolite response elements (*cre*) to cause carbon catabolite repression or activation of gene transcription.

The mechanism of carbon catabolite control has been extensively studied in *B. subtilis*. Fujita (2009) has composed an extensive review of the topic, and has compiled a list of the 50 genes experimentally identified as being under carbon catabolite control in *B. subtilis*, and their *cre* sites. Analysis of the identified *cre* sequences lead to the proposal of two consensus sequences, either with or without a TG-CA palindromic pair, WTGNAANCGNWWNCA and WTGAAARCGYTTWNN, respectively (where W is A or T, N is any

nucleotide, R is G or A, and Y is T or C). The mode of catabolite control, activation or repression was inferred from the location of the *cre* site in relation to the translation initiation base of the closest gene in its target operon (Fujita 2009). It was suggested that the binding of the activated CcpA complex to a *cre* overlapping the promoter, as for *acs* and the *acuABC* operon, would interfere with the binding of transcription machinery and result in CCR of that gene or operon. The binding of the protein complex to a *cre* site located downstream of the transcription initiation base would block transcription elongation as for 28 of the operons identified, which would also result in CCR. However, the binding of the activated CcpA complex to a *cre* upstream of the promoter, as for *ackA*, *pta*, and *ilvB*, results in carbon catabolite activation (CCA) (Moir-Blais *et al.* 2001; Presecan-Siedel *et al.* 1999). The *ackA* and *pta* genes encode acetate kinase and phosphotransacetylase, respectively, in *B. subtilis*. The *ilvB* gene is in an operon involved in the biosynthesis of branched-chain amino acids (isoleucine, valine, and leucine). In addition, the *alsSD* operon, which is involved in acetoin synthesis, is known to be positively regulated by CcpA. Thus, the CcpA-dependent CCA of these genes appears to have a role in reducing intracellular concentrations of pyruvate during growth on rich PTS-sugars by enhancing the pathways involved in excreting acetate, branched-chain amino acids, and acetoin (Figure 1.6).

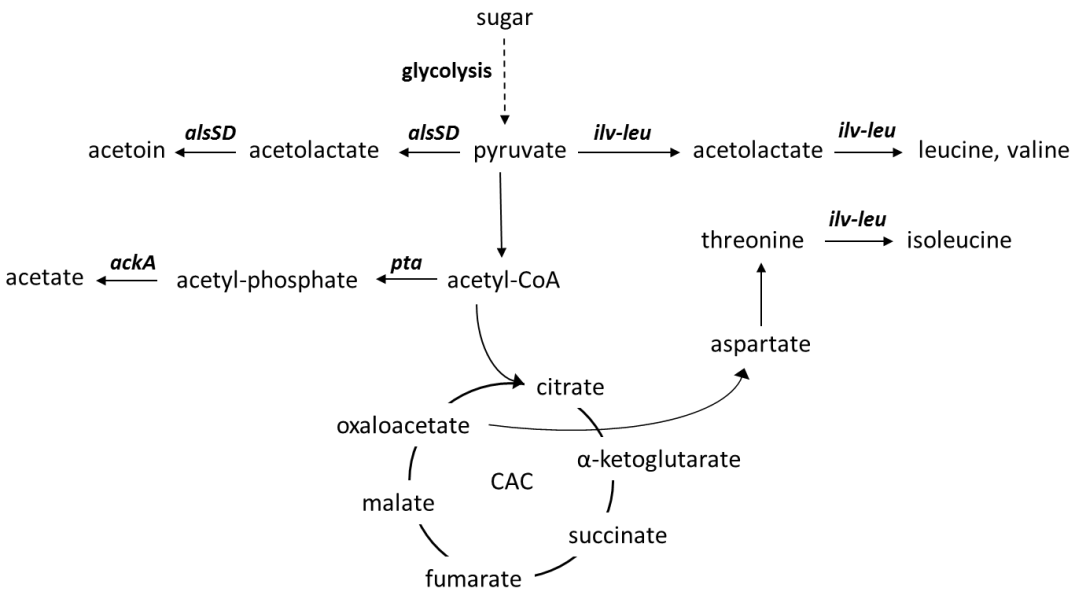


Figure 1.6 Schematic representation of carbon catabolite activated networks in *Bacillus subtilis*. The genes and operons subject to carbon catabolite activation (CCA) include the *alsSD* operon, *pta*, *ackA*, and the *ilv-leu* operon; these genes are involved in acetoin, acetate and branched-chain amino acid production.

1.7 Reducing acetate production

Acetate is often viewed as an undesirable by-product because it can retard cellular growth and inhibit protein formation. Moreover, acetate production represents a diversion of carbon that might otherwise have generated biomass or a biofuel or chemical. In prokaryotes, the major pathway of acetate production from pyruvate is via a pathway involving pyruvate formate lyase (PFL) and/or the pyruvate dehydrogenase complex (PDH), and phosphotransacetylase (PTA) and acetate kinase (AK) (Figure 1.1). A considerable amount of work has been done on reducing acetate accumulation in microbes. Approaches have included both metabolic engineering strategies and process modifications. This project focuses on metabolic engineering strategies used to disrupt acetate production.

1.7.1 Directly disrupting acetate production pathways

Deletion of genes encoding PTA and/or AK has been a common strategy to reduce acetate production; *pta* and/or *ak* have been disrupted in a variety of microorganisms including *E. coli* (Dittrich *et al.* 2005; Yang *et al.* 1999), *Clostridium acetobutylicum* (Green *et al.* 1996; Kuit *et al.* 2012; Sillers *et al.* 2008), *Salmonella enterica* (Starai *et al.* 2005), and *Bacillus subtilis* (Presecan-Siedel *et al.* 1999). The PTA and/or AK-deficient strains often have reduced acetate accumulation compared to the parent wild-type strain, but does not result in the elimination of acetate accumulation, suggesting alternative acetate-producing pathways exist. It has been observed that the *ak*-deleted mutants exhibit a reduced growth rate compared to the parent strain, particularly when grown on a rich medium. In general, the *pta* mutants do not exhibit this behaviour; therefore it has been proposed that the growth defect observed from the *ak* mutants is due to the accumulation of acetyl-phosphate.

It has been suggested that acetyl-phosphate is involved in global regulatory events due to its ability to donate its phosphoryl group to certain members of the family of two-component signal transduction (2CST) systems (Wolfe 2005; 2010). Acetyl-phosphate, in addition to other small phosphor-donors, and ATP plus histidine kinase (HK), can phosphorylate the response regulator (RR) of the 2CST system leading to the RRs activation (Figure 1.7). These activated RRs are known to control a host of cellular functions including flagella rotation, nitrogen assimilation, phosphate assimilation, expression of outer membrane porins, and biofilm development (Wolfe 2005). Thus, acetyl-phosphate is implicated in being a global regulatory molecule.

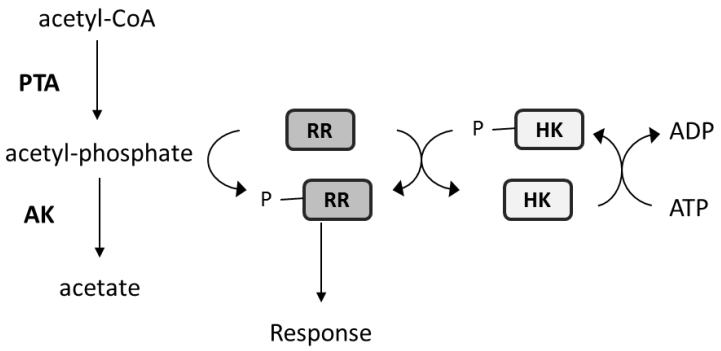


Figure 1.7 Regulation of two-component signal transduction systems by phosphoryl group donation from acetyl-phosphate. **R**esponse regulators (RRs) are activated by phosphorylation from either histidine kinases (HKs) or small molecular phosphor-donors such as acetyl-phosphate. Activated RRs are implicated in regulating a wide range of cellular functions.

1.7.2 Alternative acetate generating pathways

To the author's knowledge, elimination of acetate production by disruption to *pta* and/or *ak* has not been reported, suggesting alternative acetate-producing pathway(s) must exist in the organisms studied. Alternative suggestions have included pyruvate oxidase (PoxB), which catalyses the decarboxylation of pyruvate to form acetate and CO_2. However, disruption of *poxB* from an *E. coli* strain already deficient of PTA did not alter acetate production further, suggesting PoxB was not involved in producing acetate in this strain, and that another pathway is responsible (Phue *et al.* 2010). Another study reported a *Salmonella enterica* strain with a disrupted *ald6* gene produced less acetate than the wild-type strain, suggesting the aldehyde dehydrogenase (Ald6) is involved in acetate production (Eglinton *et al.* 2002). Other reports suggest acetate is generated via the butanediol cycle from *B. subtilis* (Speck and Freese 1973); however, disruptions to both the PTA-AK pathway and the butanediol cycle through a *pta*/*alsS* double mutant still resulted in acetate accumulation, suggesting another pathway must exist (Presecan-Siedel *et al.* 1999). In addition, a butyrate kinase (BK) has been suggested to compensate for the loss of AK in a *Clostridium acetobutylicum* strain with a disruption to *ak*, based on the substrate promiscuity of the BK enzyme (Kuit *et al.* 2012).

1.7.3 Alternative approaches to reduce acetate production

In addition to directly disrupting the pathways involved in acetate production, other genetic approaches have been used based on manipulating the central carbon metabolism by altering the phosphotransferase system (PTS), glycolysis or the pyruvate branch point (De Mey *et al.* 2007). Elimination of acetate production has been successful by deleting the PTS-

encoding genes from *E. coli*; however, the maximum growth rate was reduced by >40% (Chen *et al.* 1997; Ponce 1999). Alternatively, an approach involving the use of a harmonised pyruvate decarboxylase and aldehyde dehydrogenase pathway, which converts pyruvate to ethanol via an acetaldehyde intermediate, has been attempted (Van Zyl *et al.* 2014). However, to date an engineered organism with the ability to produce high yields of ethanol and zero acetate has not been reported.

1.8 Aims and objectives

At the onset of this project, the *Geobacillus thermoglucosidasius* TM242 strain was the current production strain utilised by TMO Renewables Ltd. This strain has been reported to produce ethanol in yields in excess of 90% of the theoretical maximum, and generate small but significant amounts of acetate (Cripps *et al.* 2009). The project aimed to both study the complex myriad of factors that control metabolic flux from acetyl-CoA to acetate in *G. thermoglucosidasius* TM242, and to investigate potential strategies of reducing the accumulation of this undesirable organic acid by-product by further metabolic manipulation.

The activities of phosphotransacetylase (PTA), acetate kinase (AK) and acetyl-CoA synthetase (ACS) are involved in acetate metabolism. The project aimed to study the activity of these enzymes by determining the kinetic properties from native cell lysates and from recombinant proteins. This would involve the cloning of *pta*, *ak* and *acs* genes from *G. thermoglucosidasius* TM242 and the characterisation of purified recombinant proteins. In addition, factors potentially capable of regulating the *in vivo* activities of these enzymes would be studied, including those that alter enzyme concentration and modify activity by post-translational modification. This in connection with the study characterising aldehyde dehydrogenase/alcohol dehydrogenase (ADHE) by a previous PhD student, Dr Jonathan Extance (2012), would be used to model and understand the regulation of metabolic flux from acetyl-CoA in this organism.

The project also aimed to investigate potential strategies of reducing the undesirable organic acid by-product, acetate. Approaches were to include the deletion or disruption to *pta* and/or *ak* genes, or alternatively the down-regulation of gene expression in order to minimise acetate production from the *G. thermoglucosidasius* TM242 strain. The resulting strain would be analysed to determine whether enhanced bioethanol production was feasible using metabolic engineering.

1.9 Project outline

Chapter 3 was a joint investigation between the author and a previous PhD student, Dr Jonathan Extance; the activities of pyruvate dehydrogenase (PDH), phosphotransacetylase (PTA), acetate kinase (AK), and aldehyde dehydrogenase/alcohol dehydrogenase (ADHE) were monitored from cell pellets collected throughout a batch-fermentation of *G. thermoglucosidasius* TM242 on a glucose-based medium. The data suggest that changes in enzyme expression and/or enzyme activity during the stages of fermentation are significant factors that determine the flux from acetyl-CoA, particularly to ethanol. In addition, the kinetic parameters of PTA, AK, and ADHE were reported from native cell lysates.

The characterisation of the recombinant PTA and AK enzymes were described in **Chapter 4** and **Chapter 5**, respectively. The purified His-tagged enzymes were characterised in terms of their kinetic and thermal properties. The studies conducted with the recombinant PTA enzyme identified an initial miss-annotation of the gene, where it was determined that the true translation initiation site was downstream of the site originally identified. Comparison of PTA and AK activity to that of ADHE suggested that the carbon flux had a greater potential towards acetate generation and not ethanol. However, it is known that ethanol is the major fermentation product, indicating that additional factors must regulate the flux from acetyl-CoA.

Chapter 6 describes how the transformation efficiency of *G. thermoglucosidasius* with the integration plasmid, pTMO31, was improved by exchanging the kanamycin resistance cassette for one that expresses a thermostable variant. This improved integration plasmid was used to delete *pta* from *G. thermoglucosidasius* TM242 and the strain was characterised after growth on both glucose- and xylose-containing media. The resulting strain had significantly reduced PTA activity, but acetate production had not been eliminated, suggesting the presence of an alternative pathway.

The *G. thermoglucosidasius* strain lacking *pta* retained around 5% of the PTA-specific activity, compared to the parent TM242 strain. It was hypothesised that the remaining PTA-specific activity could be the result of an additional phosphotransacylase (PTAC) that could catalyse the conversion of acetyl-CoA and orthophosphate to acetyl-phosphate and CoA in the Δ*pta* strain, and hence could be involved in producing acetate. **Chapter 7** describes the

cloning and characterisation of two additional PTAC enzymes, phosphotranspropionylase (PTP) and phosphotransbutyrylase (PTB). Both enzymes were discovered to accept acetyl-CoA as a substrate, but a gene deletion strategy indicated that PTB was the most likely enzyme to be involved in acetate generation from the Δ*pta* strain. The combined deletion of *pta*, *ptp* and *ptb* however, resulted in a strain that still generated acetate, suggesting that an alternative acetate-producing pathway exists in *G. thermoglucosidasius*.

Chapter 8 describes the study of acetyl-CoA synthetase (ACS) from *G. thermoglucosidasius* TM242. Four genes were annotated from the genome as encoding an ACS, and the work in this Chapter confirmed that two genes encoded proteins with ACS function. The regulatory mechanism of lysine acetylation was also investigated with ACS, including the function of a lysine acetyltransferase and a lysine deacetylase. In addition, the regulation of PTA by acetylation was investigated; the results suggest a potential regulatory response mechanism to changes in environmental conditions.

Chapter 9 is a general conclusion that summarises the major findings of the Thesis and discusses the complex myriad of factors that regulate metabolism, specifically involved at the acetyl-CoA branch-point of fermentation. The potential strategy of disrupting the PTA-AK pathway in order to reduce acetate production and enhance bioethanol yield is analysed here.

Chapter 2

Materials and Methods

2.1 Materials

All chemicals were purchased from Sigma-Aldrich (Poole, UK), unless otherwise stated. Tryptone, yeast extract, nutrient broth and agar were all purchased from Melford (Suffolk, UK), unless otherwise stated.

2.2 Bacterial strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table 2.1.

Table 2.1 Strains and plasmids used in this study.

Strain or plasmid	Description	Source or reference
E. coli strains		
JM109	A *recA*$^-$ strain also lacking the K restriction system, which prevents restriction of cloned DNA and recombination with host chromosomal DNA, improving yield and quality of isolated plasmid DNA.	Promega
BL21(DE3)	A strain deficient in proteases Lon and OmpT that uses the T7 expression system for protein expression.	NEB
G. thermoglucosidasius strains		
TM242	High ethanol producing strain with disruption to *ldh* and *pfl*, and up-regulated expression of *pdh*.	(Cripps *et al.* 2009)
TM444	TM242 background with the ability to sporulate prevented by disruption to *spo0A*.	TMO Renewables
Δ*pta*	A *pta* deletion in TM242 background.	This study
Δ*ptp*	A *ptp* deletion in TM242 background.	This study
Δ*ptb*	A *ptb* deletion in TM242 background.	This study
Δ*pta*Δ*ptp*	A *pta* and *ptp* double deletion in TM242 background.	This study
Δ*pta*Δ*ptb*	A *pta* and *ptb* double deletion in TM242 background.	This study
Δ*ptp*Δ*ptb*	A *ptp* and *ptb* double deletion in TM242 background.	This study
Δ*pta*Δ*ptp*Δ*ptb*	A *pta*, *ptp* and *ptb* triple deletion in TM242 background.	This study
Plasmids		
pGEM®-T easy	General purpose *E. coli* ligation and cloning vector incorporating TA cloning.	Promega
pET-vectors	*E. coli* protein expression vector using the T7 promoter and carries optional His-tag.	Merck Millipore
pUC19	General purpose *E. coli* ligation and cloning vector.	NEB
pTMO31	Vector designed to manipulate *G. thermoglucosidasius* chromosomal DNA, it contains pUB110 rep and kan^R genes from pUB190 in a truncated pUC19 backbone.	(Cripps *et al.* 2009)
pUCG3.8	Thermostable *Geobacillus* spp.-*E. coli* shuttle vector containing kan^R gene from pBST22.	(Bartosiak-Jentys *et al.* 2013)
pUB31	The thermostable variant of kan^R from pUCG3.8 was used to replace the mesophilic kan^R in pTMO31.	This study

2.3 Bacterial growth media

Media components were dissolved in distilled water and were sterilised by either autoclaving at 121°C for 20 min, or using 0.22µm Steritop® filter units (Merck Millipore, Darmstadt, Germany). Solid plates of the various liquid media were made by adding 1.5% (w/v) agar prior to autoclaving. Where required, media were supplemented with antibiotics at the following concentrations: 30µg/ml kanamycin (12.5µg/ml for *Geobacillus* strains carrying plasmids with kanamycin resistance markers), 50µg/ml carbenicillin. All media used for *G. thermoglucosidasius* cultures were pre-warmed to 60°C prior to inoculation.

2.3.1 Lysogeny Broth (LB)

LB consisted of 1% (w/v) NaCl (Fisher Scientific, Loughborough, UK), 1% (w/v) tryptone and 0.5% (w/v) yeast extract.

2.3.2 Yeast Extract Nutrient Broth (YENB)

YENB consisted of 0.75% (w/v) yeast extract and 0.8% (w/v) nutrient broth.

2.3.3 Super Optimal broth with Catabolite repression (SOC)

SOC consisted of 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10mM NaCl and 2.5mM KCl. This solution was autoclaved and allowed to cool before adding 100× concentrated forms of filter sterilised magnesium and glucose solutions to final concentrations of 10mM $MgCl_2$, 10mM $MgSO_4$, and 20mM glucose.

2.3.4 Soy Peptone Yeast Extract (No Glycerol) (2SPYNG)

2SPYNG consisted of 1.6% (w/v) soy peptone (Solabia, Paris, France), 1% (w/v) yeast extract (Oxoid, Basingstoke, UK), and 0.5% (w/v) NaCl, adjusted to pH 7.0 with 10M KOH. 2SPY medium is 2SPYNG with the addition of 1.0% (w/v) glycerol.

2.3.5 Tryptone Soya Broth (TS)

TS Broth (Oxoid, Basingstoke, UK) was purchased pre-prepared as a dehydrated medium containing glucose that was dissolved as recommended, 30g/L. This results in a final glucose concentration of 2.5% (w/v).

2.3.6 Modified Two Tryptone Yeast Extract (TTY)

TTY consisted of 1.6% (w/v) tryptone (Oxoid, Basingstoke, UK), 1% (w/v) yeast extract (Oxoid, Basingstoke, UK), and 0.5% (w/v) NaCl, adjusted to pH 7.0 with 10M KOH.

2.3.7 Glycerol Seed Culture

The growth medium consisted of 0.33M glycerol (Melford, Suffolk, UK), 1% (w/v) yeast extract (Oxoid, Basingstoke, UK), 25mM NaH_2PO_4, 50mM urea (Melford, Suffolk, UK), 25mM K_2SO_4, 5mM citric acid, 3.125mM $MgSO_4$, 50µM $CaCl_2$, 2.5µM biotin, and 12.5ml/L of trace elements, pH 7.0. The medium was filter sterilised to prevent the Maillard reaction occurring.

2.3.8 Modified Urea Salts Medium (USM)

USM consisted of 0.5% (w/v) yeast extract (Oxoid, Basingstoke, UK), 25mM NaH_2PO_4, 50mM urea, 25mM K_2SO_4, 5mM citric acid, 3.125mM $MgSO_4$, 50µM $CaCl_2$, 2.5µM Na_2MoO_4, and 12.5ml/L of trace elements, pH 6.7. Various concentrations of the appropriate sugar were added as indicated and the volume was brought up to 880ml prior to filter sterilisation. The medium was buffered by adding 40ml of each pre-sterilised buffer (stock concentration 1M and at pH 7.0) to bring the final concentration to 40mM Bis-Tris, 40mM HEPES, and 40mM MOPS; 2.5µM biotin was added also.

2.3.9 Trace Elements

The trace elements solution consisted of 60mM H_2SO_4, 0.144% (w/v) $ZnSO_4.7H_2O$, 0.556% (w/v) $FeSO_4.7H_2O$, 0.169% (w/v) $MnSO_4.H_2O$, 0.025% (w/v) $CuSO_4.5H_2O$, 0.0562% (w/v) $CoSO_4.7H_2O$, 0.006% (w/v) H_3BO_3, and 0.0886% (w/v) $NiSO_4.6H_2O$, dissolved in MilliQ water.

2.4 Glycerol stocks

Strains were stored long-term in glycerol suspensions at -80°C in cryogenic vials. *E. coli* glycerol stocks were made by mixing healthy cultures and sterile glycerol to a final concentration of 20% (v/v) glycerol. *G. thermoglucosidasius* glycerol stocks were made by mixing healthy cultures and sterile 50% (w/v) glycerol to a final concentration of 25% (v/v) glycerol. Glycerol stocks were snap-frozen in liquid nitrogen and immediately transferred to storage at -80°C.

2.5 *G. thermoglucosidasius* growth conditions

2.5.1 Tube fermentation of *G. thermoglucosidasius*

Characterisation of *G. thermoglucosidasius* strains was initially carried out in 15ml falcon tubes. Sterile USM (10ml) containing the desired sugar was inoculated with 1ml of an aerobically-grown seed culture that was in the exponential phase. Cultures were incubated at 60°C for greater than 48h in a shaking incubator at 250rpm. This allows the growth to transition from aerobic to micro-aerobic conditions as the oxygen becomes consumed. Supernatants were analysed by HPLC and cell lysate was assayed for enzymic activity where indicated. Unless otherwise stated, strains were analysed from three separate cultures and results are given as an average.

2.5.2 Bench-top fermenter growth of *G. thermoglucosidasius*

Batch growth in BIOSTAT® fermenters (Sartorius-Stedium, Surrey, UK) was either carried out at Bath or at TMO Renewables Ltd (see Table 2.2). The custom fermenter suite at TMO Renewables Ltd was equipped with a Tandem CO_2 and O_2 gas analyser (Magellan Instruments, Middlesex, UK). Assembly, sterilisation and operation of the fermenters followed the appropriate protocol. Seed cultures of *G. thermoglucosidasius* strains were grown in the required medium, usually either glycerol seed medium or 2SPYNG, and were grown at 60°C in a shaking incubator. Healthy cultures would typically reach an OD_{600} value of between 3 and 8.

The fermenter was inoculated with 10% (v/v) of the seed culture, and cells were grown on USM medium containing the appropriate sugar, yeast extract concentration and with or without an amino acid supplement (non-disclosed, TMO Renewables). Cultures were maintained at 60°C and at pH 6.7 by the addition of 5M KOH or 5M H_3PO_4 as appropriate. Foam levels were controlled by the addition of Antifoam 204. Temperature, pH, oxygen and foam levels were monitored throughout. Samples were taken at regular intervals to monitor the OD at 600nm, and for supernatant analysis by high-performance liquid chromatography (HPLC). Once an OD_{600} of approximately 5 was reached, the stirrer speed and aeration were changed according to the parameters in Table 2.2 to encourage transition to micro-aerobic conditions. Where necessary, cells were harvested by centrifugation, typically at 3,200 × g for 30 min in a Beckman Coulter Allegra™ 25R Centrifuge.

Table 2.2 Agitation and aeration parameters of typical fermentations used in this report. Cultures were grown initially using the pre-anaerobic switch conditions. Once an OD_{600} reading of approximately 5 was reached, the aeration and agitation (stirrer speed) were changed to the post-anaerobic switch parameters to encourage micro-aerobic conditions, and hence fermentation.

		University of Bath	TMO Renewables	
		2.5L BIOSTAT®B	10L BIOSTAT® CT-DCU	1.6L BIOSTAT® B Plus
Typical culture volume (L)		2.0	3.0	1.0
Pre-anaerobic switch	Aeration (L/min)	1.0	3.0	1.0
	Agitation (rpm)	600	440	600
Post-anaerobic switch	Aeration (L/min)	1.0	0.6	0.2
	Agitation (rpm)	600	220	300

2.5.3 Analysis of fermentation products

Fermentation products were quantified in clarified culture samples by HPLC, either at Bath or at TMO Renewables Ltd. Both sites used a Hewlett Packard 1100 system, and detection was by UV absorption at 215nm for organic acids, and by refractive index for sugars and ethanol using an Agilent Technologies 1200 series LC system (at Bath) or a Knauer WellChrom K-2301 RI detector (at TMO). The column used at Bath was an Aminex® HPX-87H ion exclusion column (BioRad) (typical retention times of analytes studied are stated in Table 2.3), while TMO Renewables Ltd used a Nucleogel Sugar 810H column (Machegey Nagel). The flow rate was 0.6ml/min, and temperature was set to 30°C. The mobile phase was 5mM sulphuric acid, with a 5µl injection and a run time of 27min.

Table 2.3 Typical retention times of analytes studied in this thesis on the Aminex® HPX-87H ion exclusion.

	Retention time (min)	
Molecule	RID	UV (215nm)
Glucose	8.9	
Xylose	9.6	
Pyruvate		9.3
Lactate		12.6
Acetate		15.4
Ethanol	21.4	

2.6 Molecular biology

2.6.1 *G. thermoglucosidasius* genomic DNA extraction

G. thermoglucosidasius genomic DNA was prepared following a TMO Renewables extraction protocol originally adapted from a Novacta method (Novacta Biosystems Ltd, Herts, UK). Cells were revived from glycerol stocks stored at -80°C by inoculating 10ml of TTY medium with a loop full of cells. Cells were grown overnight at 52°C in a shaking incubator set to 250rpm. Cells were pelleted by centrifugation at 4,500 × g for 20min in a Jouan B4i Centrifuge and resuspended in 1ml of lysis buffer (10-500µg/ml RNase A and 10mg/ml lysozyme, dissolved in autoclaved sterilised buffer; 0.3M sucrose (Fisher Scientific, Loughborough, UK), 25mM Tris-HCl (Fisher Scientific, Loughborough, UK), 25mM EDTA (Melford, Suffolk, UK) pH 8.0), before being transferred to 1.5ml tubes. Tubes were incubated at 30°C for 1h, before addition of 5µl of proteinase K (20mg/ml proteinase K, 50mM Tris-HCl pH 8.0, 1.5mM calcium acetate). SDS was then added to a final concentration of 0.1% (w/v) before incubation at 50°C for 1h.

The sample was transferred to a 2ml centrifuge tube before addition of an equal volume of P:C:I (phenol:chloroform:isoamyl alcohol). The sample was gently mixed before centrifugation at 12,000 × g for 10min. The top aqueous layer was transferred to a new 2ml centrifuge tube, before an equal volume of C:I (chloroform:isoamyl alcohol) was added and mixed as before. The sample was centrifuged at 12,000 × g for 10min before the translucent top layer was transferred to a new tube. 3M sodium acetate, pH 6.0, was added to give a final concentration of 300mM before the sample was mixed well. 0.7 volume of isopropanol was added, and the sample was mixed gently. The sample was centrifuged at 12,000 × g for 10min before the supernatant was removed. The pellet was washed with 70% ethanol and centrifuged at 12,000 × g for 10min. This ethanol wash was repeated before the pellet was allowed to dry in a fume hood. The pellet was resuspended in 50µl of MilliQ water and the DNA was stored at -20°C.

2.6.2 Polymerase chain reaction

All genes were PCR-amplified from TM242 genomic DNA using Phusion®HF (Finnzymes, Loughborough, UK), following the manufacturer's instructions. PCR-amplification was carried out in 20µl reactions using an Eppendorf Mastercycler gradient PCR thermocycler (Eppendorf, Cambridge, UK). A typical reaction contained 1× Phusion HF buffer, 0.2mM

dNTPs, 0.25µM of each appropriate primer, and variable concentrations of template DNA. The reaction was started by the addition of 0.4U of Phusion Hot start DNA polymerase, before a three-step amplification protocol was followed.

2.6.3 Agarose gel electrophoresis

DNA was visualised and purified by electrophoresis in a 1% (w/v) agarose gel, with 0.5µg/ml ethidium bromide. Agarose was dissolved in TAE buffer (40mM Tris-HCl, 57.1% (v/v) glacial acetate (Fisher Scientific, Loughborough, UK), and 1mM EDTA, pH 8.0). DNA samples were mixed with 6× DNA loading dye (0.25% (w/v) Bromophenol blue (BDH Prolabo, Leicestershire, UK), 0.25% (w/v) Orange G, 50mM EDTA and 30% (v/v) glycerol). Samples were run in parallel to a 1kb ladder (Bench Top 1kb DNA ladder, Promega, Southampton, UK) before being visualised on a UV transilluminator. DNA was purified using the Promega Wizard SV gel and PCR clean-up system following the manufacturer's protocol.

2.6.4 pGEM®-T easy vector cloning

PCR-amplified DNA fragments were A-tailed using *Taq* DNA polymerase. A typical 10µl reaction contained 70% (v/v) purified PCR product, 1× *Taq* buffer, 0.2mM dATP and 5U of *Taq* DNA polymerase. The mixture was incubated at 70°C for 30min, and the A-tailed product was purified using the Wizard SV gel and PCR clean-up Kit. A-tailed PCR products were ligated into the linear pGEM®-T easy vector using the 3'- terminal thymine. The use of this vector allowed additional screening of transformants by using the blue/white screening method.

2.6.5 Ethanol precipitation

Ethanol precipitation was used to purify ligation products. To the ligation products 10% (v/v) 3M sodium acetate, pH 5.2, 10% (v/v) dextran blue and 200% (v/v) of ice-cold 100% ethanol were added. After mixing, reactions were stored at -20°C for >30min before centrifugation at 13,000 × g for 20min at 4°C. After removal of the supernatant, ice-cold 70% (v/v) ethanol was added, and the sample spun again. This was repeated twice before the pellet was allowed to dry in a fume hood. The pellet was resuspended in MilliQ water to 50% of original volume.

2.6.6 Preparation of electrocompetent *E. coli*

E. coli JM109 cells were revived from glycerol stocks by growing overnight in YENB at 37°C in a shaking incubator set to 200rpm. YENB in baffled flasks was inoculated with 1% (v/v) of the overnight culture, and cells were grown until mid-exponential phase. Cells were harvested by centrifugation at 3,400 × g for 10min at 4°C before being resuspended in ice-cold MilliQ water. Cells were centrifuged and washed with ice-cold MilliQ water a further three-times before final resuspension in 10% (v/v) sterile glycerol to 0.3% of the original culture volume. 50µl aliquots were snap-frozen on dry ice and stored at -80°C.

2.6.7 Transformation of electrocompetent *E. coli*

Thawed *E. coli* JM109 cells were incubated with 0.5µl of required plasmid for >2min. Aliquots were then transferred to Genepulser™ cuvettes of 1mm pathlength, before transformation using a Micropulser™ electroporator (Bio-Rad, Hertfordshire, UK), at 1.8kV and typical time constants of 5ms. Transformed cells were grown in 1ml of SOC for 1h at 37°C in a shaking incubator. Between 50-200µl of transformants were spread onto LB plates containing appropriate antibiotics. LB plates for pGEM-T transformants had 100µl of 100mM IPTG (Melford, Suffolk, UK) and 20µl of 50mg/ml X-Gal (Melford, Suffolk, UK) spread onto the plates to allow blue/white selection.

2.6.8 Transformation of chemically competent *E. coli*

E. coli BL21(DE3) cells were transformed with purified pET vectors by heat shock. 20µl of cells were incubated on ice with 1µl of plasmid DNA for >20min. Samples were then heat shocked at 42°C for 45s in a heat block or water bath, before being incubated on ice for >2min. 1ml of SOC medium was added and cells were incubated at 37°C for 1h before 50µl of transformants were spread on LB plates containing appropriate antibiotics. Plates were incubated at 37°C overnight to allow colony growth.

2.6.9 Preparation of electrocompetent *G. thermoglucosidasius*

G. thermoglucosidasius cells were revived from glycerol stocks by growing overnight in 2SPYNG at 60°C in a shaking incubator set to 220rpm. Cells from 1ml of this culture were harvested by centrifugation at 3,400 × g and used to inoculate 50ml of 2SPYNG in a sterile 250ml baffled flask. Cultures were grown at 60°C in a shaking incubator set to 220rpm until an OD_{600} of 1.4 was reached. Cultures were cooled on ice for 10min prior to centrifugation

at 3,400 × g for 20min at 4°C. The supernatant was discarded and cell pellets were resuspended in ice-cold electroporation medium containing 10% (w/v) glycerol, 0.5M sorbitol, and 0.5M mannitol. Cells were centrifuged and washed with ice-cold electroporation medium a further four-times before final resuspension in electroporation medium to 2% of the original culture volume.

2.6.10 Transformation of electrocompetent *G. thermoglucosidasius*

Aliquots (60µl) of electrocompetent *G. thermoglucosidasius* cells were incubated on ice with the desired plasmid. Aliquots were then transferred to Genepulser™ cuvettes of 1mm pathlength, before transformation using a Micropulser™ electroporator at 2.5kV, 25µFD, 600Ω and typical time constants of 5ms. Transformed cells were recovered in 1ml of 2SPY for 2h at between 55°C and 60°C (or 4h at 52°C if using pTMO31 based constructs) in a shaking incubator. Cells were harvested by centrifugation at 1,800 × g for 2min. The majority of the supernatant was discarded before cells were resuspended in approximately 100µl of supernatant before spreading out onto TS agar plates containing 12.5µg/ml kanamycin.

2.6.11 Blue/white colony screening

Transformants with plasmids containing the *lacZα* gene (pGEM®-T, pUC19, pTMO31 and pUB31) were screened using the blue/white protocol. The *lacZα* gene is disrupted upon successful ligation of a target gene into the multiple cloning sites of these vectors. Disruption of the *lacZα* gene prevents β-galactosidase activity in transformants and is observed as white colonies on agar plates containing 50mg/ml X-gal and 100mM IPTG. Transformants of unsuccessful ligation products would be observed as blue colonies.

2.6.12 Colony PCR screen

Colony PCR reactions were performed to screen large numbers of colonies that were produced from transformations with newly-ligated vectors. A typical reaction (20µl) contained 1× Taq Master Mix (GeneSys Ltd, Surrey, UK) [*Taq* DNA polymerase, Taq buffer, dNTPs and loading dye], and 0.25µM of each primer (either gene or vector specific). The DNA was introduced to the mixture by addition of a few cells from each colony, which were lysed during the initial denaturation step. The PCR protocol in Table 2.4 was followed, and

steps 2-3 were cycled 25-30 times. Successful transformants were identified by PCR products of the expected gene size.

Table 2.4 The colony PCR protocol with *Taq* DNA polymerase. The elongation time is dependent on the size of the amplicon (1min/kb).

Step	Temperature (°C)	Time
1 – Initial denaturation	95	30 s
2 – Denaturation	95	30 s
3 – Annealing	55	30 s
4 – Elongation	68	X min
5 – Final elongation	68	10 min

2.6.13 Purification of plasmid DNA

E. coli JM109 strains carrying plasmids were cultured overnight at 37°C in a shaking incubator at 200rpm, with appropriate antibiotic. The cells were centrifuged at 3,400 × g for 10min and the supernatant removed. Plasmid DNA was purified using the Promega Wizard plus SV miniprep DNA purification system following the manufacturer's protocol.

2.6.14 Restriction digest

DNA samples were digested with the appropriate restriction enzyme following the manufacturer's protocol. Most restriction enzymes were either from New England BioLabs® Inc (NEB) or Promega. Double digests were performed in the manufacturer's suggested buffer, and where appropriate bovine serum albumin (BSA) was added. Where appropriate, digestion was halted using heat-treatment and shrimp alkaline phosphatase (SAP) was used to inhibit self-ligation of vectors.

2.6.15 Sequencing

The success of the cloning was verified by sending the plasmid DNA to be sequenced by Source BioScience LifeSciences (London, UK) or Eurofins Genomics (Ebersberg, Germany). T7F and SP6 primers were used for pGEM-T plasmids, and T7F and T7R primers for pET vectors. Other vector or gene specific primers were used where required. Sequencing results were analysed by the GeneDoc programme and the BioEdit Sequence Alignment Editor Software package.

2.6.16 Creation and isolation of gene knockouts in *G. thermoglucosidasius*

Electrocompetent *G. thermoglucosidasius* cells were transformed with the desired construct as described in Section 2.6.10. Successful transformants were grown in 10ml of 2SPYNG in a 50ml falcon tube at 60°C in a shaking incubator for 6-8h. Cells were harvested by centrifugation at 3,400 × g for 4min. Harvested cells were resuspended in approximately 100µl of supernatant and spread onto TS agar plates containing 12.5µg/ml kanamycin. These plates were incubated at 68°C for 16h, which forces the integration of the plasmid into the chromosome due to the thermolabile replication of origin from pUB110.

Primary integrants were selected and grown in 2SPYNG at 55°C - 60°C in a shaking incubator. Cultures were then passaged up to eight times. After each subculture, cells were diluted and spread onto TS agar plates (without antibiotic) and grown at 60°C. Successful colonies were replica-plated onto TS agar plates containing 12.5µg/ml kanamycin. Kanamycin sensitive colonies were screened by colony PCR to determine whether the strain had reverted back to parental genotype or the knockout genotype had been isolated.

2.7 Protein biochemistry

2.7.1 Protein expression

Recombinant proteins were expressed from pET vectors in *E. coli* BL21(DE3) cells using the T7 RNA polymerase expression system. Transformants were selected for overnight culture in 10ml of LB with appropriate antibiotic. 1% (v/v) of overnight culture was used to inoculate fresh LB in baffled conical flasks, in the presence of antibiotic. Cultures were grown at 37°C in a shaking incubator until an OD_{600} of 0.8-1.0 was reached. Protein expression was induced for 4.5h by the addition of IPTG to a final concentration of 1mM. Cells were pelleted at 3,200 × g for 10min at 4°C and stored at -20°C until required.

2.7.2 *E. coli* cell extract preparation

E. coli BL21(DE3) cell pellets were re-suspended to one-tenth of the original culture volume in His-bind buffer (50mM Tris-HCl, pH 8.0, 300mM NaCl, and 20mM imidazole (Acros Organics, Geel, Belgium)). One complete Mini, EDTA-free protease inhibitor cocktail tablet (Roche Diagnostics GmbH, Mannheim, Germany) was added per 10ml sample volume. Cells were lysed on ice by sonication at 14microns with three 20-second bursts with a 30-second cooling period between each burst. The samples were centrifuged at 16,000 × g for 20min

to remove insoluble debris. The supernatant was filtered through 0.45µm and 0.22µm Millex® Syringe driven filter units (Merck Millipore, Darmstadt, Germany) before loading onto the chromatography column.

2.7.3 *G. thermoglucosidasius* cell extract preparation

G. thermoglucosidasius cell pellets were resuspended to a concentration of 0.3g cells/ml buffer (50mM EPPS, pH 8.0). Pefabloc® SC AEBSF and Benzonase® Nuclease were added to 1mM and 350U/ml of cell suspension, respectively. The cells were lysed by sonication at 14microns on ice with three 20-second bursts with a 30-second cooling period between each burst. The suspension was centrifuged at 16,000 × g for 20 min to remove insoluble debris. The supernatant was collected and diluted in 50mM EPPS buffer, pH 8.0, as required.

2.7.4 Nickel-affinity chromatography

Nickel-affinity chromatography was used to purify recombinant proteins that had a poly(His)-tag. Metal-chelating cellulose (Bioline, London, UK) was loaded into a Poly-Prep chromatography column (Bio-Rad Laboratories GmbH, Munich, Germany) and charged with Ni^{2+} by passing 2× column volumes (CV) of 400mM $NiSO_4$ through the column. The column was equilibrated with His-bind buffer prior to loading of the soluble protein fraction. Bound proteins were eluted by sequential washes with increasing imidazole concentration in a 50mM Tris-HCl, pH 8.0, buffer containing 300mM NaCl and 20-1000mM imidazole. Imidazole was removed by dialysis prior to enzyme assays, unless otherwise stated.

2.7.5 Protein estimation

Protein concentrations were determined by the Bradford method (Bradford 1976). Reaction mixtures (1ml) contained 20% (v/v) Bio-Rad Protein Assay Dye Reagent (Bio-Rad Laboratories GmbH, Munich, Germany) and the protein sample. Samples were incubated at room temperature for 20 min before the A_{595} was recorded. Standard curves (linear with typical R^2 values >0.95) were produced using dilutions of bovine serum albumin (Pierce, Rockford, USA), such that protein concentrations were in the range 0-10µg/ml. Protein concentration estimates of enzyme samples were determined in this linear range.

2.7.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE analysis was performed using a Mini-PROTEAN electrophoresis system (BioRad Laboratories GmbH, Munich, Germany), using gels made in-house. Gels (10%) were produced using the components described in Table 2.5. Ammonium persulphate (APS) and Tetramethylethylenediamine (TEMED) were added immediately prior to casting, to catalyse polymerisation of the gel. Samples were prepared by heating at 98°C for 3min with 4× SDS-PAGE loading buffer containing 0.2M Tris-HCl, pH 6.8, 2% (w/v) SDS, 40% (v/v) glycerol, 0.1% (v/v) β-mercaptoethanol (VWR, Leicestershire, UK) and 0.8mg/ml Bromophenol Blue. Gels were run at 400mA and 200V in 50mM MOPS, 50mM Tris-HCl, pH 7.7, buffer containing 1mM Ethylenediaminetetraacetic acid (EDTA) and 0.1% (w/v) SDS. Protein bands were stained by washing in a solution containing 0.25% (w/v) Coomassie Brilliant Blue R-250, 45% (v/v) methanol, and 10% (v/v) acetic acid. Protein bands were visualised by washing in destaining solution containing 20% (v/v) methanol and 10% (v/v) acetic acid.

Table 2.5 SDS-PAGE gel components.

Component	Volume of component added (µl)	
	Stacking gel	Resolving gel
1.25M *bis*-Tris-HCl, pH 6.6	1000	2840
30% (w/v) acrylamide	460	3340
MilliQ water	2004	3820
10% (w/v) APS	40	50
TEMED	20	20

2.7.7 Measurement of enzyme activities

Enzyme activities were measured at 60°C, unless otherwise stated, using a Varian Cary 50 or Cary 60 Bio UV/visible light spectrophotometer and a Peltier temperature controller (Agilent Technologies, Cheshire, UK). Specific activities in U/mg protein (µmol/min/mg) were determined using the molar absorption coefficient of the detected molecule, and the protein concentration of the measured sample. Where detected, substrate-independent background rates were subtracted from enzyme rates recorded. Substrates were either dissolved in the assay buffer or as recommended by the supplier. It was confirmed that under all assay conditions the rate of reaction was directly proportional to the enzyme concentration.

2.7.7.1 Acyl-CoA preparation

To obtain 1ml of an approximately 7mM concentration of acyl-CoA solution, 10mg of coenzyme-A (CoA) were dissolved in 1ml of MilliQ water, and the solution was cooled on ice for 10 min. Then, 0.2ml of 1M potassium bicarbonate (pH 8.0) was added, after which 5µl of the required acyl-anhydride (Acros organics, Geel, Belgium) was added and the solution was incubated on ice for at least a further 10 min. Complete acylation was confirmed when no increase in absorbance at 412nm was observed when a sample of the acyl-CoA solution was mixed with 5,5'-dithiobis(2-nitro-benzoic acid) (DTNB). DTNB is cleaved by thiol compounds stoichiometrically to release a yellow compound, 2-nitro-5-thiobenzoate^{2-} (NTB^{2-}) (Ellman 1959). The concentration of NTB^{2-} can be quantified by the A_{412nm} (ε_{412} = 13,600 $M^{-1}.cm^{-1}$).

2.7.7.2 Determination of acetyl-CoA concentration

The stock concentration of acetyl-CoA used in the enzymatic assays was determined spectrophotometrically. This was achieved by measuring the increase in A_{412nm} (ε_{412} = 13,600 $M^{-1}.cm^{-1}$) corresponding to the release of CoA by citrate synthase activity in the presence of excess oxaloacetate. Citrate synthase catalyses the conversion of acetyl-CoA and oxaloacetate to citrate and CoA, a reaction that goes to >99% completion. The reaction mixture (1ml) contained 19mM Tris-HCl, pH 8.0, 0.2mM DTNB, and 10µl of acetyl-CoA. The spectrophotometer was calibrated with a 1ml cuvette loaded with this reaction mixture. 5U of citrate synthase was added to the reaction mixture, and the reaction proceeded to completion before the A_{412nm} was measured. Citrate synthase does not catalyse the reaction to the same extent with other acyl-CoAs, and therefore the concentration of these was presumed to be 7mM.

2.7.7.3 Phosphotransacylase activity

$$\text{acyl-CoA} + \text{phosphate} \rightleftarrows \text{acyl-phosphate} + \text{CoA}$$

Phosphotransacylase (PTAC) activity was measured by following the increase in A_{412nm} (ε_{412} = 13,600 $M^{-1}.cm^{-1}$), corresponding to the released CoA reacting with DTNB. Typical reaction mixtures (1ml) contained 48mM Hepes, pH 6.5 at 60°C, 0-120mM $Na_2HPO_4.12H_2O$ (Fisher Scientific, Loughborough, UK), 0.1mM DTNB, 0.0-0.7mM acetyl-CoA and enzyme. It is important to note that addition of cell extract/enzyme is used to start the reaction because DTNB inactivates phosphotransacetylase (PTA) in the absence of acyl-CoA (Dr Tracy Tillotson, personal communication, University of Bath).

2.7.7.4 Acetate kinase

acetyl-phosphate + ADP ⇄ acetate + ATP

Kinase activity was measured in the organic acid-forming direction by determining the rate of ATP formation in a coupled enzymatic assay with hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6P-DH) by the method of Bowman et al (1976) (Figure 2.1). The reaction mixture (1ml) contained 50mM Hepes, pH 6.5 at 60°C, 5mM $MgCl_2.6H_2O$, 2.5mM D-(+)-glucose, 0.4mM $NADP^+$ (Melford, Suffolk, UK), 0.0-8.0mM ADP, 0.0-5.0mM acetyl-phosphate, 1.2U of G6P-DH, and 2U of HK. The reaction was started by adding cell extract/enzyme, and the rate of NADPH formation was measured at 340nm (ε_{340} = 6,200 $M^{-1}.cm^{-1}$).

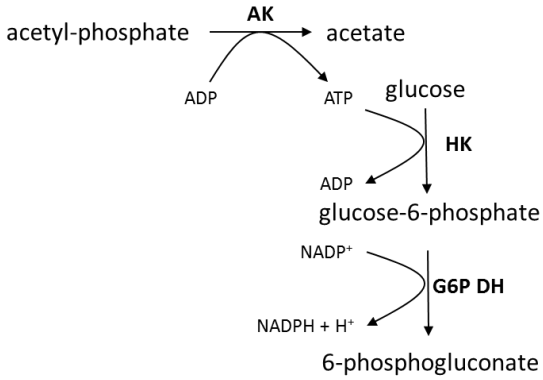


Figure 2.1 Schematic representation of the acetate kinase coupled assay. The ATP production by acetate kinase (AK) was coupled to the activities of hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6P DH), and the rate of NADPH production was measured at 340nm.

2.7.7.5 Determining kinetic parameters

Kinetic parameters were determined using the enzyme kinetics module of SigmaPlot (Systat Software Inc, London, UK). Where enzyme activities obeyed Michaelis-Menten kinetics, kinetic parameters were determined using non-linear regression based on the Michaelis-Menten equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

Some of the enzyme activities observed in this study appeared to have been influenced by substrate inhibition. Kinetic parameters were determined from these data by fitting the data to the substrate inhibition equation:

$$v = \frac{V_{max}[S]}{K_m + [S] + \left(\frac{[S]^2}{K_i}\right)}$$

Some of the enzyme activities observed in this study appeared to have been influenced by positive cooperative substrate binding. Kinetic parameters were determined from these data by fitting the data to the Hill equation:

$$v = \frac{V_{max}[S]^n}{{K_m}^n + [S]^n}$$

Chapter 3
Enzyme activities in native cell lysates

3.1 Introduction

The *Geobacillus thermoglucosidasius* TM242 strain with disruptions to genes encoding lactate dehydrogenase (LDH) and pyruvate formate lyase (PFL), and with up-regulated expression of the pyruvate dehydrogenase complex (PDH), has been reported to produce ethanol as the major product of fermentation (Cripps *et al.* 2009). However, small but significant quantities of acetate are also generated during fermentation by this strain. Considering the application of *G. thermoglucosidasius* TM242 in industrial bioethanol production, acetate is an undesirable by-product because production of this organic acid decreases the pH of the medium, thereby inhibiting growth of the culture. Moreover, acetate production represents a diversion of carbon that may have alternatively generated ethanol.

Compared to the wild-type strain, *G. thermoglucosidasius* TM242 has increased metabolic flux from sugar to acetyl-CoA (Figure 1.1). The conversion of acetyl-CoA to ethanol is catalysed by the bifunctional aldehyde dehydrogenase/alcohol dehydrogenase (ADHE) enzyme. The phosphotransacetylase (PTA) and acetate kinase (AK) form a pathway which catalyses the conversion of acetyl-CoA to acetate via an acetyl-phosphate intermediate. The 'competition' for acetyl-CoA between ADHE and PTA will have significant effects on the ratio of ethanol to acetate formed during fermentation.

A wide-variety of factors will be involved in controlling the flux from acetyl-CoA in *G. thermoglucosidasius* TM242, including those that regulate the intracellular concentration of ADHE, PTA and AK, and the activities of these enzymes. The concentration of an enzyme is dependent on transcription, the level of transcripts, translation and protein degradation. The ability of an enzyme to catalyse a reaction is dependent on enzyme concentration, kinetic properties, post-translational modification, and intracellular substrate concentrations. To date, these factors have not been studied for either the PTA or AK from *G. thermoglucosidasius*.

The *pta* and *ak* genes are located together in an operon in most bacteria, including *E. coli*, indicating that these two genes are cotranscribed (Brown *et al.* 1972; 1977; Latimer and

Ferry 1993). Alternatively, in other bacteria including *G. thermoglucosidasius* and *Bacillus subtilis* the *pta* and *ak* genes are located at distinct loci on the chromosome (Presecan-Siedel *et al.* 1999). The distinct location of the *pta* and *ak* genes indicates that an alternative mechanism must control the expression of these two genes. It has been reported that the expression of both *pta* and *ak* in *B. subtilis* is controlled by carbon catabolite regulation (Grundy *et al.* 1993a; Presecan-Siedel *et al.* 1999).

As reviewed in Section 1.6, carbon catabolite control is a mechanism by which bacteria can adapt to environmental stimuli and selectively metabolise a preferred carbon source in the presence of more than one carbohydrate. The uptake of a preferred sugar, such as glucose, leads to an increase in fructose-1,6-bisphosphate, which stimulates HPr kinase. HPr kinase phosphorylates HPr and Crh at Ser-46, which can then bind to CcpA. This activated CcpA complex can then bind to catabolite response elements (*cre*), resulting in the carbon catabolite repression or activation of the gene or operon it controls (Figure 1.5).

Analysis of the identified *B. subtilis cre* sequences lead to the proposal of two consensus sequences, either with or without a TG-CA palindromic pair, WTGNAANCGNWWNCA and WTGAAARCGYTTWNN, respectively (where W is A or T, N is any nucleotide, R is G or A, and Y is T or C) (Fujita 2009). It has been reported that the proximity of the *cre* site in relation to the transcription initiation base dictates the mode of catabolite response in *B. subtilis* (gene repression or activation) (Fujita 2009). The binding of the activated CcpA complex to a *cre* site upstream of the promoter activates transcription of that gene/operon. This is the case for *pta* and *ak* in *B. subtilis*, where the metabolism of a PTS-preferred sugar such as glucose activates the transcription of these genes (Grundy *et al.* 1993a; Presecan-Siedel *et al.* 1999).

Interestingly, regions that closely resemble or exactly match the *cre* consensus sequences identified from *B. subtilis* can be found 5'- of the predicted promoter regions for both *pta* and *ak* in *G. thermoglucosidasius* (Figure 3.1). This suggests a possible role for carbon catabolite control of these genes from *G. thermoglucosidasius* TM242. The implication of the location of these putative *cre* sites is yet to be studied, however.

(A) WTGNAANCGNWWNCA – Sequence 1 WTGAAARCGYTTWNN – Sequence 2

(B) ATGTAAAAGATAATAAACTGAATTGTCTTAATAAATCGATTTTTTATGGCAATGCTTAACTTAGTTGC
GTTTCTATATAATAAGAGTATCCTAGTATATAGGGGCATTTCAAACATTAGGGAGGATTCGTTCGTG

(C) TTGTAAACGATTACATATAGAAATTGCTTGAAAGAAAACATAAACAATCATACAATGAAAAAAGATAT
GTGGAATAGTGAAAAAATGAACAAAATATGACAAAAGATG

Figure 3.1 Location of the putative *cre* sequences for *G. thermoglucosidasius* TM242 genes. Regions upstream of the *pta* (B) and *ak* (C) initiator codon (highlighted green) were analysed to identify putative *cre* sequences and promoter sites. The grey highlighted regions represent nucleotides that are similar to the *B. subtilis cre* sequences described by Fujita (2009) (A). Yellow and blue highlighted regions represent the -10 and -35 boxes, respectively, of the promoters identified using the Softberry BPROM online promoter prediction tool.

To date, the parameters controlling the flux from acetyl-CoA to acetate in *G. thermoglucosidasius* have not been studied. This Chapter describes the investigation of ADHE, PTA and AK from native cell lysates. The kinetic properties of these enzymes determined from the same cell lysate, under the same conditions, are reported here. In addition, the relative activity of PDH, ADHE, PTA and AK from cell lysate samples collected during a batch-fermentation of *G. thermoglucosidasius* TM444 on a glucose-based medium are reported in this Chapter. Thus, the relative changes in expression and/or enzyme activity can be determined and related to the production of fermentation metabolites. This, in addition to the kinetic properties determined, will help to generate a metabolic model of carbon flux from acetyl-CoA.

The studies in this Chapter were completed in collaboration with a previous PhD student, Dr Jonathan Extance, who completed the work on ADHE and PDH (Extance 2012). The *G. thermoglucosidasius* TM444 (a sporulation-deficient strain based on a *spo0A* deletion in the TM242 background) fermentation and analysis were completed by the team at TMO Renewables.

3.2 Methods

3.2.1 *G. thermoglucosidasius* growth

The *G. thermoglucosidasius* strains required for determining both kinetic parameters and monitoring of the expression/activity levels were grown in a bench-top fermenter as outlined in section 2.5.2.

The kinetic parameters were determined from *G. thermoglucosidasius* TM242 grown at the University of Bath in a USM medium containing 6% (w/v) sucrose and 0.5% (w/v) yeast

extract. The inoculum for this fermentation was from an overnight culture of *G. thermoglucosidasius* TM242 on the glycerol seed medium. Cells were harvested once the culture had depleted oxygen levels and the OD_{600} had reached 6.7.

Enzyme expression levels were determined from *G. thermoglucosidasius* TM444 cells grown at TMO Renewables on a USM medium containing 6% (w/v) glucose, 2% (w/v) yeast extract and amino acid mix 1 (non-disclosed, TMO Renewables). The inoculum for this growth was from a *G. thermoglucosidasius* TM444 culture grown on 2SPYNG that had reached an OD_{600} of 12.6. The fermentation culture volume was initially 5L, and samples were taken and analysed by the team at TMO Renewables.

3.2.2 Enzyme assays

Cell lysates were prepared as described in Section 2.7.3 and enzyme assays were performed as described in Section 2.7.7, unless otherwise stated. All activities, except PDH, were conducted in 50mM Hepes buffer, with 0.1mM zinc acetate, pH 6.5 at 60°C. The buffer used to characterise AK activity included 5mM $MgCl_2$. The ionic strength of the assay was maintained at 0.1M by the addition of NaCl.

3.2.2.1 Phosphotransacetylase activity

Kinetic parameters for PTA were determined using the protocol described in Section 2.7.7.3. PTA activity levels were measured in triplicate in *G. thermoglucosidasius* TM444 cell lysates using 50mM sodium orthophosphate, pH 7.0, and 0.12mM acetyl-CoA.

3.2.2.2 Acetate kinase activity

Kinetic parameters for AK were determined by using the protocol described in Section 2.7.7.4. AK activity levels were measured in *G. thermoglucosidasius* TM444 cell lysates using 1mM acetyl-phosphate and 1mM ADP.

3.2.2.3 Aldehyde dehydrogenase activity

$$\text{acetyl-CoA} + \text{NADH} + H^+ \rightleftarrows \text{acetaldehyde} + \text{CoA} + NAD^+$$

AldDH activity was measured by following the increase in absorbance at 412nm (ε_{412} = 13,600 $M^{-1}.cm^{-1}$), corresponding to the released CoA reacting with DTNB. Typical reaction mixtures (1ml) contained 48mM Hepes, pH 6.5 at 60°C, 0.1mM zinc acetate, 0.0-0.4mM

acetyl-CoA, 0.0-0.3mM NADH, 0.1mM DTNB and enzyme. AldDH activity levels were measured in *G. thermoglucosidasius* TM444 cell lysates using 0.35mM acetyl-CoA and 0.18mM NADH.

3.2.2.4 Alcohol dehydrogenase activity

$$\text{acetaldehyde} + \text{NADH} + H^+ \rightleftarrows \text{ethanol} + NAD^+$$

ADH activity was measured by following the decrease in absorbance at 340nm (ε_{340} = 6,200 $M^{-1}.cm^{-1}$), corresponding to the direct utilisation of NADH. ADH activity levels were measured in *G. thermoglucosidasius* TM444 cell lysate using 40mM acetaldehyde and 0.18mM NADH.

3.2.2.5 Pyruvate dehydrogenase activity

$$\text{pyruvate} + \text{CoA} + NAD^+ \rightleftarrows \text{acetyl-CoA} + CO_2 + \text{NADH} + H^+$$

PDH activity was measured by following the increase in absorbance at 340nm (ε_{340} = 6,200 $M^{-1}.cm^{-1}$), corresponding to the direct production of NADH. PDH activity levels were measured in *G. thermoglucosidasius* TM444 cell lysates in a 50mM potassium orthophosphate buffer, pH 7.0 at 60°C. The reaction mixture (1ml) contained 50mM potassium orthophosphate, pH 7.0, 2.8mM NAD^+, 0.2mM thiamine pyrophosphate (TPP), 1.0mM $MgCl_2$, 0.13mM coenzyme-A (in cysteine-HCl), 2mM sodium pyruvate and enzyme.

3.3 Results

3.3.1 Phosphotransacetylase

3.3.1.1 Effect of ionic strength

Phosphotransacetylase (PTA) catalyses the conversion of acetyl-CoA and orthophosphate to acetyl-phosphate and CoA. The production of CoA can be measured by its reaction with 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) by following the change in absorbance at 412nm. Dr Tracy Tillotson (University of Bath) previously identified that PTA from *G. thermoglucosidasius* is inactivated by DTNB, but that the enzyme was protected by the presence of substrates. Therefore, to maximise initial rates, the enzyme was the last reagent to be added to the assay.

During the initial characterisation of PTA, it was observed that PTA appeared to exhibit substrate inhibition with orthophosphate (data not shown). However, orthophosphate carries a high negative charge, which would increase the ionic strength of the assay. Therefore, to understand whether the inhibition of PTA observed was due to increasing ionic strength or not, the activity of PTA was measured using the standard assay in the presence and absence of 0.9M NaCl. The results indicated that NaCl also inactivates PTA (Figure 3.2). In addition, substrate inhibition was no longer observed when the ionic strength was maintained whilst varying orthophosphate concentrations, suggesting that the negative charge carried by orthophosphate causes the inactivation.

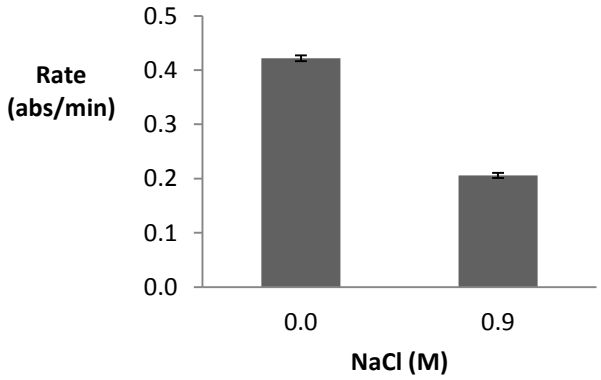


Figure 3.2 Inactivation of PTA activity by addition of NaCl. The rate of PTA activity was compared with or without the addition of 0.9M NaCl to the assay.

3.3.1.2 Native enzyme kinetics

PTA is a bisubstrate enzyme; in the acetate-forming direction, acetyl-CoA and orthophosphate are its two substrates. The rate equation for bisubstrate enzymes takes the form:

$$v = \frac{V_{\max}}{1 + \frac{K_{\mathrm{m}A}}{[A]} + \frac{K_{\mathrm{m}B}}{[B]} + \frac{K_{\mathrm{s}A}K_{\mathrm{m}B}}{[A][B]}}$$

Where v is the initial rate, $V_{\max}$ is the maximum rate, $[A]$ and $[B]$ are the concentrations of substrate A and B respectively, $K_{\mathrm{m}A}$ and $K_{\mathrm{m}B}$ are the Michaelis constants for substrate A and B respectively, and $K_{\mathrm{s}A}$ is the dissociation constant for substrate A. It is possible to obtain a good estimate of the kinetic parameters of a bisubstrate enzyme by saturating the assay with one of the substrates (i.e. where $[B] > 10\times K_{\mathrm{m}B}$), in which case the rate equation then reverts to the simpler form:

$$v = \frac{V_{\max}}{1 + \frac{K_{\mathrm{m}A}}{[A]}}$$

However, PTA is affected by changes in ionic strength, and therefore saturating orthophosphate concentrations could not be used if near physiological ionic strength (approximately 0.1M) was to be maintained in the assay. To maintain the ionic strength at 0.1M whilst varying orthophosphate concentrations, NaCl was added.

To determine accurate kinetic parameters for PTA, a matrices approach was taken. That is where apparent kinetic parameters can be determined by varying one substrate and keeping the second substrate at a fixed concentration. By repeating this at several different fixed concentrations of the second substrate, the true kinetic parameters can be determined from the range of apparent parameters.

G. thermoglucosidasius TM242 cells grown on USM medium with 6% (w/v) sucrose and 0.5% (w/v) yeast extract were lysed and the soluble supernatant was analysed for enzyme activity. PTA displayed typical Michaelis-Menten kinetics towards phosphate concentration at each fixed concentration of acetyl-CoA (Figure 3.3 A). Secondary plots were used to determine a $V_{\max}$ of 4.5 (± 0.1) U/mg and a $K_{\mathrm{m(acetyl\text{-}CoA)}}$ of 0.059 (± 0.006) mM (Figure 3.3 B and C). PTA also displayed typical Michaelis-Menten kinetics towards acetyl-CoA concentration at each fixed concentration of phosphate (Figure 3.4 A). Secondary plots were used to determine a $V_{\max}$ of 4.5 (± 0.5) U/mg and a $K_{\mathrm{m(phosphate)}}$ of 11 (± 3) mM (Figure 3.4 B and C).

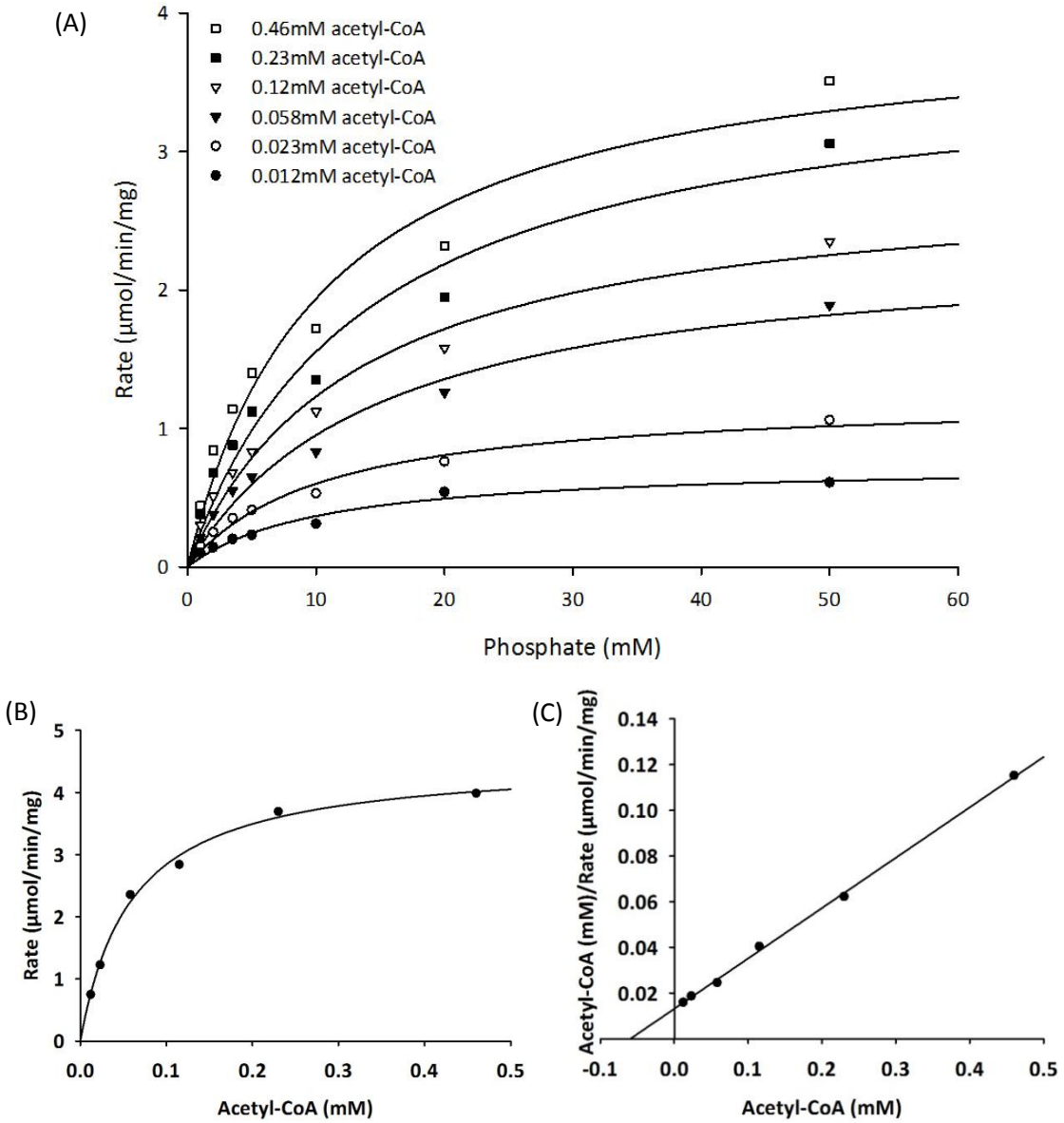


Figure 3.3 Native PTA activity from a *G. thermoglucosidasius* TM242 cell lysate. The relationship between specific activity (U/mg) and phosphate concentration at a variety of fixed acetyl-CoA concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true K_m for acetyl-CoA at saturating phosphate concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

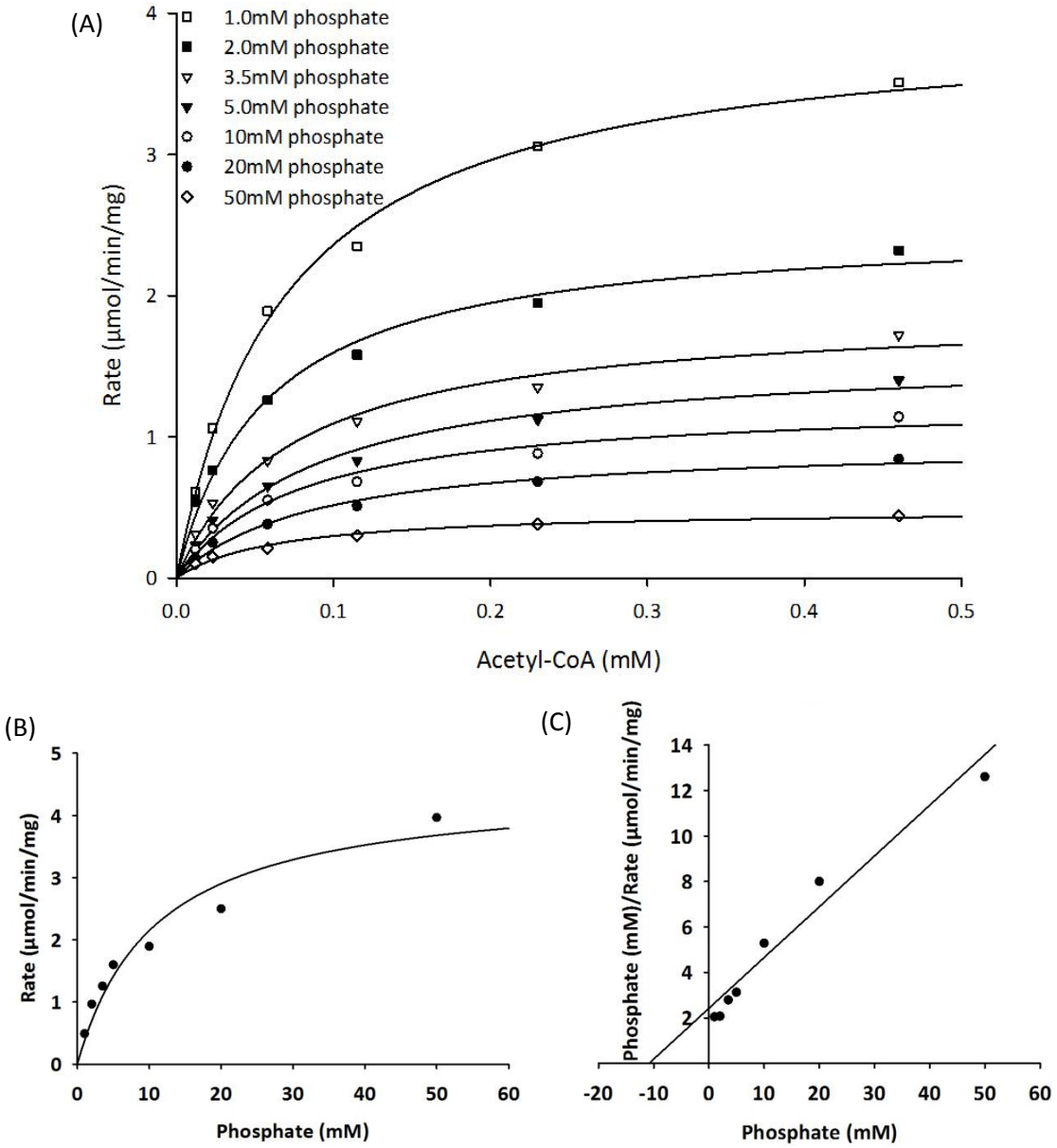


Figure 3.4 Native PTA activity from a *G. thermoglucosidasius* TM242 cell lysate. The relationship between specific activity (U/mg) and acetyl-CoA concentration at a variety of fixed phosphate concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true K_m for phosphate at saturating acetyl-CoA concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

3.3.2 Acetate kinase

3.3.2.1 Assay development

Acetate kinase activity measured in the acetate-forming direction is typically assayed by a coupled assay system, measuring the liberation of ATP using the activities of hexokinase and glucose-6-phosphate dehydrogenase (Section 2.7.7.4). The coupled enzymes used for the AK assays were from mesophilic sources, and therefore susceptible to rapid protein degradation at the elevated temperatures of the assay. However, the initial rates measured

were directly proportional to the amount of AK/TM242 cell lysate added to the assay. This suggested that the rate limiting step was the amount of AK enzyme in the assay and that the mesophilic enzymes were active and in excess (Figure 3.5).

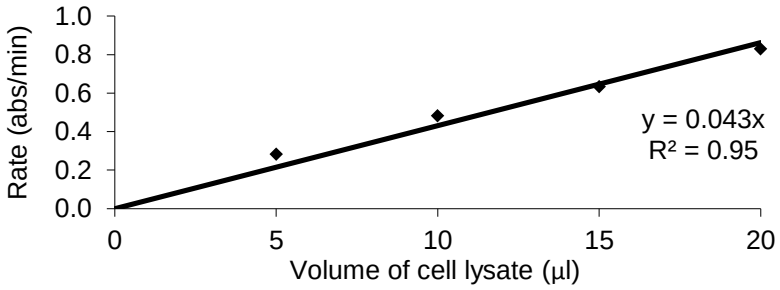


Figure 3.5 Dependence of enzyme activity on enzyme concentration.

3.3.2.2 Inhibition

It was observed that the AK enzyme appeared to display substrate inhibition with ADP. To determine whether this was an effect of changes in ionic strength similar to the PTA activity, the ionic strength of the assay was maintained at 0.1M by addition of NaCl. The substrate inhibition became less prominent but was still evident, suggesting that AK activity is inhibited by both increasing ionic strength and substrate inhibition with ADP (data not shown). Consequently, kinetic parameters in relation to ADP were determined by fitting the data to the substrate inhibition rate equation.

3.3.2.3 Native enzyme kinetics

Kinetic characterisation of AK from *G. thermoglucosidasius* TM242 cell lysates was determined at constant ionic strength. Substrate inhibition was observed with ADP; data were fitted to the substrate inhibition rate equation using SigmaPlot (Figure 3.6 A and B) and yielded a $V_{\max}$ of 7.7 (± 0.4) U/mg, where U is µmol of ATP produced per minute, and a $K_{\mathrm{m(ADP)}}$ of 0.18 (± 0.01) mM. The AK displayed typical Michaelis-Menten kinetics towards acetyl-phosphate concentration (Figure 3.6 C and D); data were analysed by SigmaPlot and yielded a $V_{\max}$ of 6.9 (± 0.1) U/mg and a $K_{\mathrm{m(acetyl\text{-}phosphate)}}$ of 0.25 (± 0.06) mM.

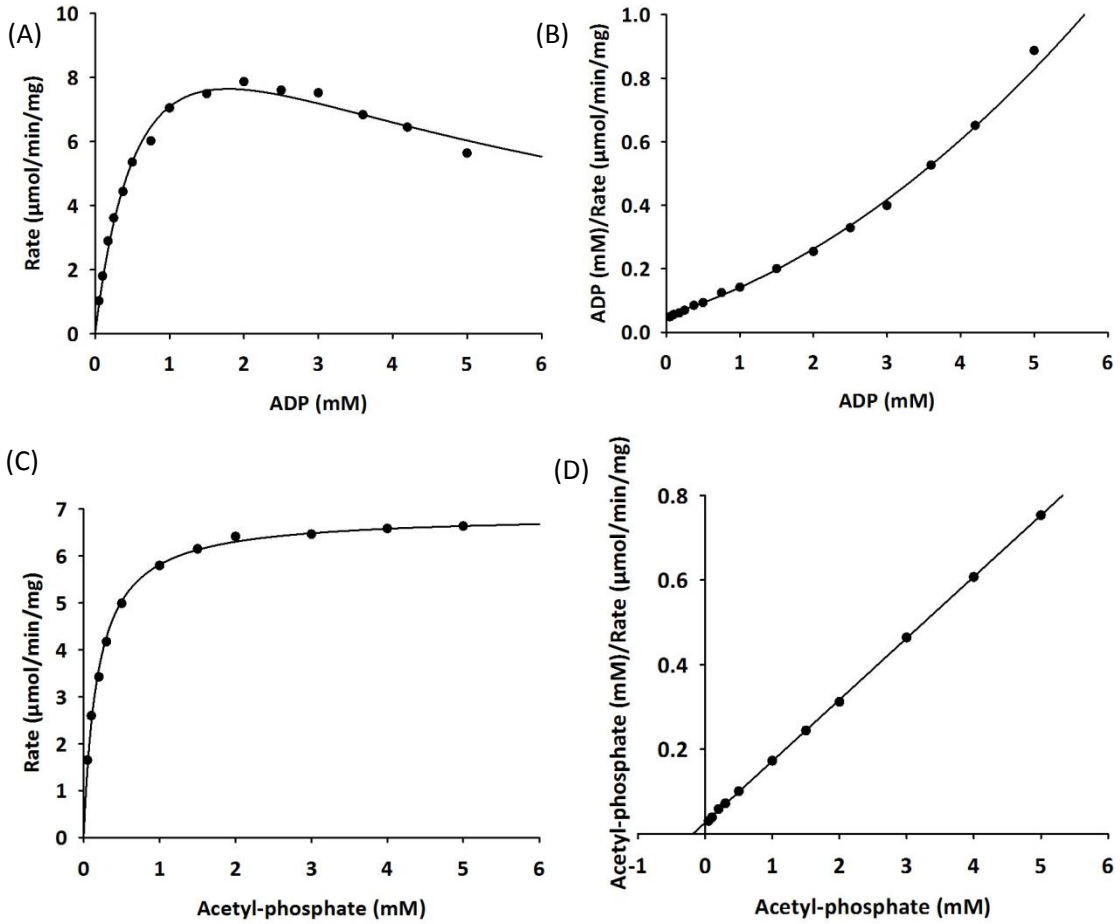


Figure 3.6 Native AK activity from a *G. thermoglucosidasius* TM242 cell lysate. The AK activity was measured as the acetyl-phosphate concentration was kept constant at 5mM while the ADP was varied; data were fitted to the substrate inhibition rate equation and are represented as (A) a Michaelis-Menten plot and (B) a Hanes-Woolf plot. In addition, the AK activity was measured as the ADP concentration was kept constant at 3mM while the acetyl-phosphate concentration was varied; data were fitted to the Michaelis-Menten rate equation and are represented as (C) a Michaelis-Menten plot and (D) a Hanes-Woolf plot.

3.3.3 Comparison of enzyme activities throughout a fermentation run

Metabolic flux is influenced by several factors, including enzyme kinetic parameters, metabolite concentrations and enzyme concentration. To determine whether expression/activity of the enzymes governing micro-aerobic acetyl-CoA flux alters during fermentation, a batch culture of *G. thermoglucosidasius* TM444 on USM medium containing 6% (w/v) glucose and 2% (w/v) yeast extract was analysed. *G. thermoglucosidasius* TM444 is an improved production strain which differs from *G. thermoglucosidasius* TM242 by disruption to genes controlling sporulation, but the fermentative ethanol pathway follows that of *G. thermoglucosidasius* TM242 (Figure 1.1). The culture was analysed at time points for OD_{600} and supernatant composition by TMO Renewables. The fermentation followed a typical profile (Figure 3.7) (personal communication, Joanne Neary, TMO Renewables Ltd).

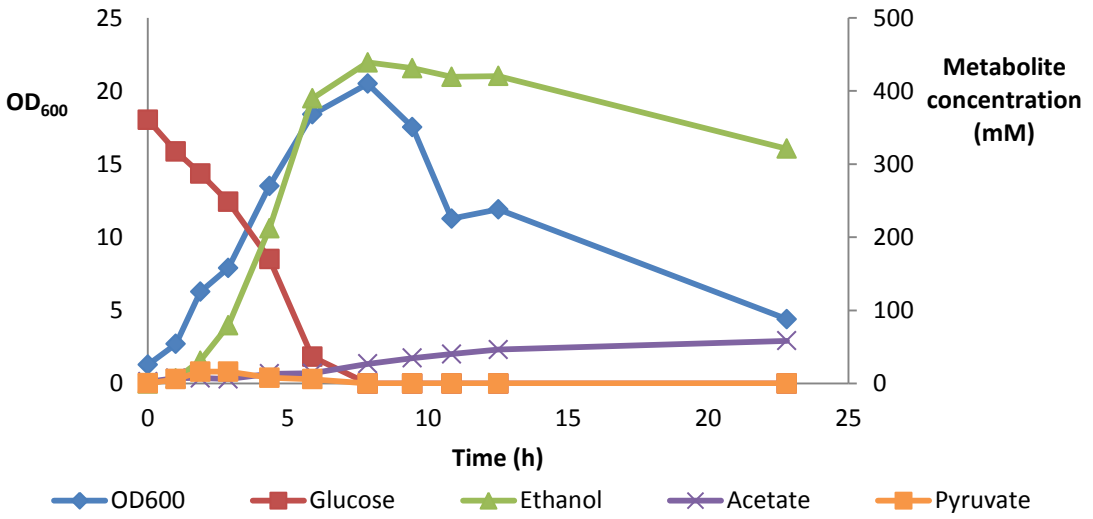


Figure 3.7 The fermentation profile of *G. thermoglucosidasius* TM444 on USM medium containing 6% (w/v) glucose and 2% (w/v) yeast extract. The supernatant was analysed at time intervals for OD_{600}, glucose consumption and product accumulation.

Cell samples collected during the fermentation were lysed and the soluble fraction was analysed for enzyme activity. PTA and ADHE activity was determined in the same 50mM Hepes buffer, pH 6.5, while the AK assay buffer had 5mM $MgCl_2$ added. The PDH enzyme was assayed according to the method developed by Dr Tracy Tillotson (University of Bath). The specific activity of PTA and AK remain higher than that of ADHE (AldDH and ADH) throughout the fermentation (Figure 3.8). However, it is worth noting that the rates recorded here are under assay conditions that may not be optimal for each enzyme nor represent physiological conditions. Therefore the change in activity relative to the maximal activity was calculated and can be observed in Figure 3.9.

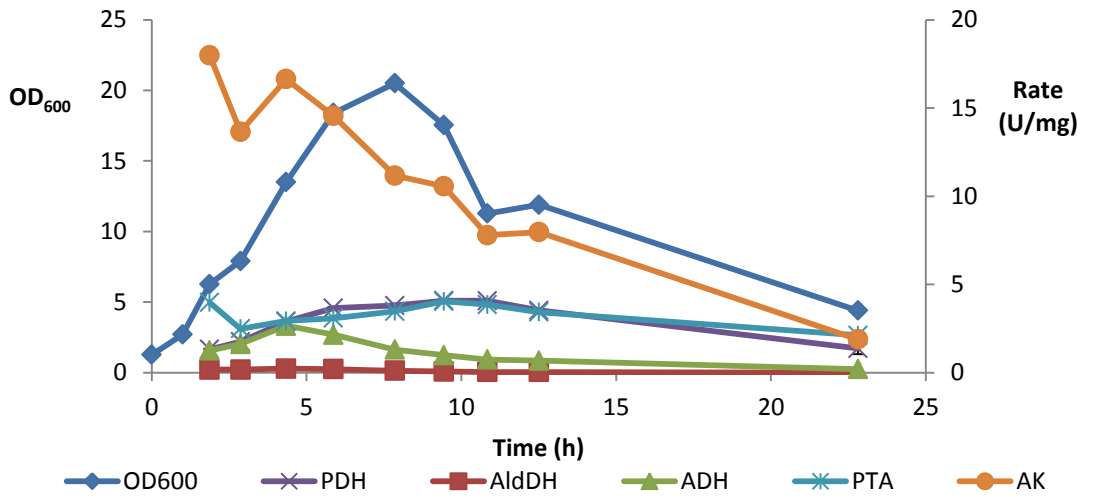


Figure 3.8 The enzyme activities governing acetyl-CoA flux from a *G. thermoglucosidasius* TM444 cell lysate. The cell pellets collected during the fermentation were lysed and enzymatic activities were measured.

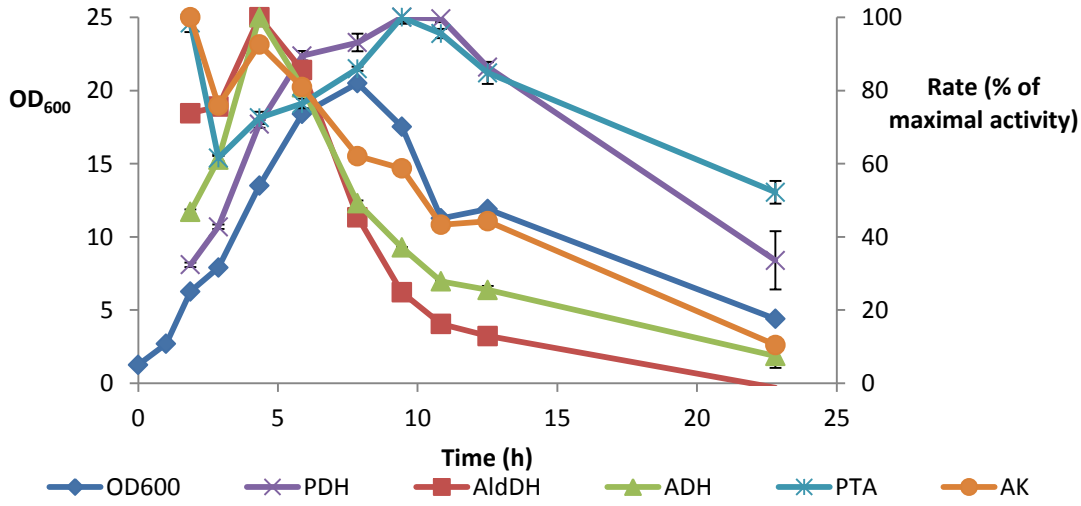


Figure 3.9 Relative enzyme activities throughout a *G. thermoglucosidasius* TM444 fermentation run. The relative activity in terms of a percentage of maximal activity measured for each enzyme was determined from the cell lysate.

The maximum rate of both AldDH and ADH activities of the bifunctional ADHE enzyme coincide with the maximum ethanol production rate (Figure 3.10). Whereas the acetate production loosely fits the trend of PTA activity, with the majority of acetate produced after the glucose has been exhausted (Figure 3.11).

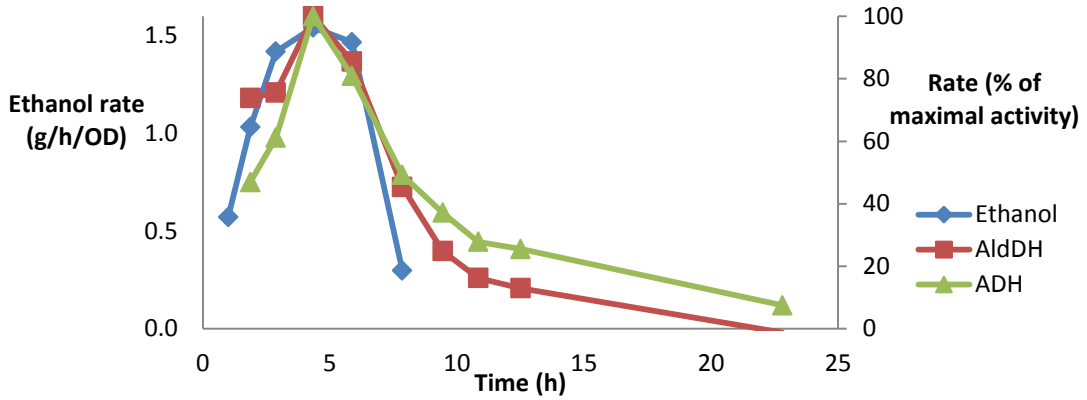


Figure 3.10 The rate of ethanol production in relation to ADHE activities. Relative rates are expressed as a percentage of maximal activity of that enzyme.

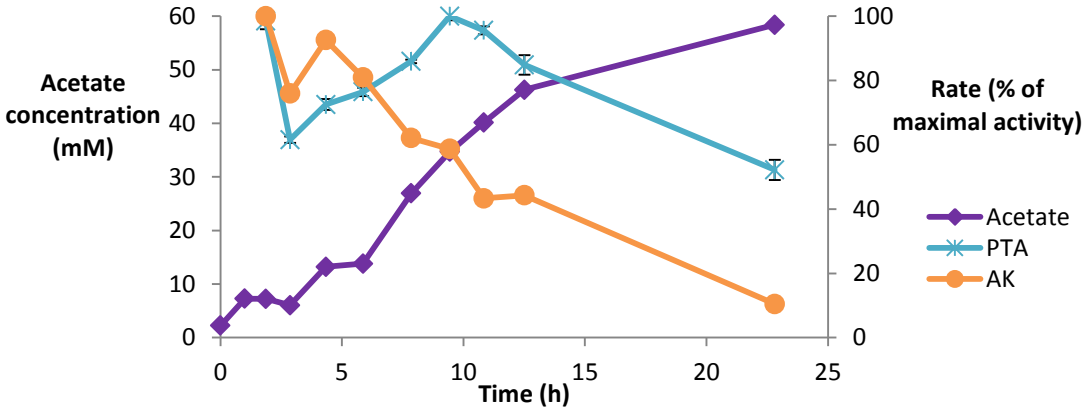


Figure 3.11 Acetate production in relation to PTA and AK activities. Relative rates are expressed as a percentage of maximal activity of that enzyme.

3.4 Discussion

Metabolic flux *in vivo* depends on enzyme concentration, kinetic properties, and other metabolic pathways that alter substrate concentrations. A qPCR technique could have been used to quantify transcript levels of the genes involved in catalysing the metabolic flux from acetyl-CoA to ethanol and acetate in *G. thermoglucosidasius* TM242. However, the concentration of an enzyme is dependent on a multitude of factors including transcription, the level of transcripts, translation and protein degradation. Moreover, the *in vivo* activity is dependent on factors that regulate enzymic properties. Therefore, the relative changes in PDH, ADHE, PTA and AK activities were investigated throughout fermentation.

The results demonstrate that the activities of PTA and AK remain higher than those of the bifunctional ADHE enzyme throughout fermentation. Even at the peak AldDH activity, the specific activity of PTA is 13-fold higher. There is, however, a limitation to comparing the

specific activities of two different enzymes characterised here because the assays utilised represent neither optimal nor *in vivo* conditions. However, it is possible to relate the expression pattern for each enzyme to product formation.

Acetyl-CoA is mostly generated by the PDH complex in *G. thermoglucosidasius* TM242. In this strain, the *pdh* operon is under the control of the anaerobic *ldh* promoter, which in comparison to the wild-type strain causes an increase in expression under micro-aerobic conditions. However, during the fermentation of *G. thermoglucosidasius* TM242, pyruvate accumulates in the medium, suggesting the PDH levels are limiting the flux through to acetyl-CoA. But once the PDH levels increase the pyruvate is taken up. This does give scope for improved pathway engineering by placing the *pdh* operon under control of a more constitutively active promoter. Of the enzymes characterised here, the PTA is the most constitutively active, and therefore its promoter may provide a better alternative.

It is evident that both the AldDH and ADH activities of the bifunctional ADHE enzyme correlate with the maximal ethanol production rate. This suggests that the production of ethanol is controlled by altering ADHE concentration. The rapid decrease in activity once the glucose is exhausted suggests the ADHE is readily degraded and that expression of this enzyme is tightly regulated. Acetate production has a weak correlation to PTA levels and is inversely correlated to AK levels; the majority is produced once the glucose has been exhausted. The activity patterns of PTA and AK suggest that alternative factors to intracellular concentrations of these enzymes may be involved in regulating the flux between acetyl-CoA and acetate. In addition, these data do not provide a conclusive answer to whether *pta* and/or *ak* expression is regulated by carbon catabolite control. A better method for this study would be to measure changes in transcript levels between hexose and pentose growth using qPCR.

To compare the activities in either the ethanol- or acetate-forming pathways competing for acetyl-CoA, the kinetic properties of PTA and AldDH were determined from the same *G. thermoglucosidasius* TM242 cell extract. The results indicate that the $V_{\max}$ value for PTA is >34-fold higher than that of AldDH (Table 3.1). The Michaelis constants of the two enzymes for acetyl-CoA are almost identical, but there is a significant difference between the K_{m} values for the other substrates; this suggests that the metabolite concentrations of orthophosphate and/or NADH may be significant factors that control metabolic flux.

Table 3.1 Kinetic parameters of *G. thermoglucosidasius* TM242 enzymes competing for acetyl-CoA.

Enzyme (substrate)	$V_{\max}$ (U/mg)	K_{m} (mM)
PTA (acetyl-CoA)	4.5 (± 0.1)	0.059 (± 0.006)
PTA (phosphate)	4.5 (± 0.5)	11 (± 3)
AldDH (acetyl-CoA)	0.094 (± 0.004)	0.052 (± 0.001)
AldDH (NADH)	0.13 (± 0.00)	0.10 (± 0.00)

The kinetic properties of ADHE have also been reported from *G. thermoglucosidasius* TM242 cell extracts in a 50mM citrate buffer, pH 6, by measuring the loss of NADH (Extance *et al.* 2013). This paper reported the $V_{\max}$ value for AldDH to be 1.2 (± 0.1) U/mg and the K_{m} value for acetyl-CoA to be 0.019 (± 0.003) mM, while the K_{m} value for NADH was 0.16 (± 0.01) mM. These kinetic parameters were determined under conditions that were more optimal for the ADHE enzyme, and results in the PTA $V_{\max}$ value being just 3.8-fold higher than that of AldDH. This suggests that the intracellular conditions, particularly the pH, can have significant effects on the enzyme's ability to catalyse metabolic flux. However, the paper suggests the $V_{\max}$ value for AldDH may be over-estimated due to the activity of ADH also catalysing the oxidation of NADH. Despite this, the *in vitro* activities determined suggest PTA has a greater potential to direct flux towards acetate when compared to the ability of AldDH to catalyse flux towards ethanol. However, the fermentation of *G. thermoglucosidasius* TM242 results in greater yields of ethanol than acetate, indicating that other factors control metabolic flux.

In summary, the results of this Chapter have confirmed that there is a higher potential flux from acetyl-CoA to acetate compared to ethanol in *G. thermoglucosidasius* TM242. However, the higher concentrations of ethanol produced from fermentation suggest that there are factors that limit the flux towards acetate production and/or stimulate the flux towards ethanol. The tight regulation of ADHE expression/activity is an apparent factor that correlates with increased ethanol production. In addition, the kinetic properties, particularly the K_{m} values of ADHE and PTA, and the metabolite concentrations may also be significant factors. The complex interplay of regulatory systems within bacteria means that it is likely that several other factors are also involved in controlling the production of ethanol and acetate in *G. thermoglucosidasius* and these are yet to be identified.

Chapter 4
Characterisation of recombinant phosphotransacetylase

4.1 Introduction

Acetyl-CoA is a pivotal metabolite in central metabolism; it is a building block for lipid and amino acid biosynthesis, cell wall synthesis, and as a precursor for the synthesis of secondary metabolites (Bentley 2000; Brown *et al.* 1977; El-Mansi 2004; Semenza 2001). During aerobic growth acetyl-CoA enters the citric acid cycle (CAC), where it is oxidised to carbon dioxide and generates the high-energy chemical ATP. The NADH generated from the CAC is reoxidised by oxidative phosphorylation using oxygen as the electron acceptor.

During growth conditions where oxygen availability is limited or when the influx of carbon into the cells exceeds the CAC capacity, ATP and NAD^+ must be generated by alternative pathways. Acetogenesis, the excretion of acetate into the environment, occurs as a result of these conditions. The conversion of acetyl-CoA to acetate is catalysed by phosphotransacetylase (PTA) and acetate kinase (AK), and generates a molecule of ATP. PTA catalyses the reversible transfer of the acetyl group between CoA and orthophosphate (Equation 4.1), and replenishes the pool of CoA. The phosphorylation of ADP by acetyl-phosphate is catalysed by AK (Equation 4.2).

$$CH_3COCoA + HPO_4^{2-} \rightleftarrows CH_3COOPO_3^{2-} + CoA \quad (4.1)$$

$$CH_3COOPO_3^{2-} + ADP^{3-} \rightleftarrows CH_3COO^- + ATP^{4-} \quad (4.2)$$

The PTA-AK pathway is primarily considered an acetate excretion pathway (Wolfe 2005), until recently thought only to exist in prokaryotes (Atteia *et al.* 2006; Ingram-Smith *et al.* 2006). Acetate assimilation occurs either by the reversal of the PTA-AK pathway or by the high-affinity system of the AMP-forming acetyl-CoA synthetase (ACS; EC 6.2.1.1).

Two classes of PTA enzymes have been discovered in microbes: class I PTAs are around 320-350 amino acids in length, and class II PTAs are around 700 amino acids in length (Alina Campos-Bermudez *et al.* 2010; Brinsmade and Escalante-Semerena 2004; 2007). Sequence alignments of PTAs from the two classes reveal that class I PTAs share homology with the C-terminal domain of class II PTAs. This suggests that the active site of the class II PTAs is located at the C-terminus of the protein. Studies on class II PTAs from *E. coli* and *Salmonella*

enterica have indicated that the N-terminus is involved with allosteric regulation from pyruvate, PEP, ADP, ATP and NADH (Alina Campos-Bermudez *et al.* 2010; Brinsmade and Escalante-Semerena 2007). Allosteric regulation of the class I PTAs has not been demonstrated (Dimou *et al.* 2011; Rado and Hoch 1973).

The *pta* and *ak* genes in most bacteria have been found to be organised into a single operon. As in *Bacillus subtilis* (Presecan-Siedel *et al.* 1999), *G. thermoglucosidasius* TM242 *pta* and *ak* genes are located at distinct loci on the chromosome. Annotation of the genome sequence was carried out for TMO Renewables Ltd by ERGO™ Integrated Genomics. The putative *pta* gene identified encodes a protein 353 amino acids in length that is 74% identical and 81% similar to the PTA from *B. subtilis*, indicating that the enzyme belongs to the class I family.

The PTA enzyme has not been characterised from *G. thermoglucosidasius*. Therefore, there is no evidence that this *pta*-like gene actually encodes a PTA. This chapter reports the characterisation of the recombinant PTA from *G. thermoglucosidasius* TM242 in terms of its kinetic and thermal properties. Cloning of *pta* into an *E. coli* expression vector revealed an additional translational start site, 3'- of the original identified by ERGO™. Mass spectrometry was used to determine the form of PTA that is native to *G. thermoglucosidasius* TM242.

Temperature optima and thermostability studies were carried out by supervised undergraduate project students, Kirsten Dundas and Hristina Ivanova. Hristina also helped develop the native PTA purification protocol. Mass spectrometry (MS) in Section 4.3.3 was carried out by Dr Anneka Lubben (University of Bath), and MS in Section 4.3.6 by Dr Catherine Botting (University of St. Andrews).

4.2 Methods

4.2.1 Cloning

Cloning of the *pta* constructs used general molecular biology techniques described in Section 2.6. Vector-specific primers were used for DNA sequencing to determine the correct sequence had been cloned. pET-28a(+) clones were transformed into *E. coli* BL21(DE3) cells for protein expression.

4.2.1.1 *pta*-pET-28a(+)

The *pta* gene identified by ERGO™ Integrated Genomics was amplified by the polymerase chain reaction (PCR) from *G. thermoglucosidasius* TM242 genomic DNA, using primers *pta*F1 and *pta*R1 (Table 4.1). The PCR product was purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *pta* gene was excised from the pGEM®-T vector by restriction digest with *Nde*I and *Xho*I, agarose gel purified and ligated into pET-28a(+) that had been linearized with *Nde*I and *Xho*I. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the construct.

4.2.1.2 Full-length *pta*-pET-28a(+)

To create a construct to express the full-length *pta*, the GTG codon was changed to GTT by site-directed mutagenesis, using a QuickChange II site-directed mutagenesis kit (Agilent Technologies, UK) (for the location of the GTG codon see the full-length phosphotransacetylase sequence in the Appendix). The manufacturer's protocol was followed using Phusion DNA polymerase with primers Full-length *pta*F and Full-length *pta*R (Table 4.1). The parental DNA was digested by treating the sample with *Dpn*I, which degrades the methylated and hemimethylated DNA. The plasmid sample was purified prior to transformation of *E. coli* JM109 cells.

4.2.1.3 Truncated *pta*-pET-28a(+)

To create a construct to express the truncated *pta*, the sequence encoding the truncated PTA was PCR-amplified from the *pta*-pET-28a(+) construct, using the primers Truncated *pta*F and *pta*R1 (Table 4.1). The PCR product was purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The truncated *pta* gene was excised from the pGEM®-T vector by restriction digest with *Nde*I and *Xho*I, agarose gel purified and ligated into pET-28a(+) that had been linearized with *Nde*I and *Xho*I. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the construct.

Table 4.1 PCR primers used for *pta* cloning. Restriction sites are underlined, and the GTG to GTT mutation is highlighted in grey.

Name	Sequence (5'-3')
*pta*F1	CATATGATGGCAATGCTTAACTTAGTTGCGTTTCT
*pta*R1	CCTCGAGGTTATAGCGATTGCGCCGCA
Full-length *pta*F	AGGGAGGATTCGTTCGTTAGCAGTGATTTAT
Full-length *pta*R	ATAAATCACTGCTAACGAACGAATCCTCCCT
Truncated *pta*F	CATATGATGAGCAGTGATTTATTTTCGACAT

4.2.2 Protein purification

Recombinant proteins were purified by nickel-affinity chromatography as described in Section 2.7.4. Purified recombinant proteins were unstable after diluting to a concentration necessary for enzymatic assays (see Section 4.3.5.1). Therefore, protein samples were diluted in buffers containing bovine serum albumin (5mg/ml).

4.2.2.1 Size-exclusion chromatography

Size-exclusion chromatography was used to purify and analyse a sample of nickel-affinity purified PTA. This technique was carried out on a Superdex™ 200 10/300 GL gel filtration column on an ÄKTA explorer FPLC system (GE Healthcare, Buckinghamshire, UK). The column was calibrated using standard proteins of various molecular weights: carbonic anhydrase, Mr: 29,000; albumin, Mr: 66,000; alcohol dehydrogenase, Mr: 150,000; β-amylase, Mr: 200,000; apoferritin, Mr: 443,000; and thyroglobulin, Mr: 669,000. The column was equilibrated in a 50mM Tris-HCl buffer, pH 8.0, with 150mM NaCl. A dialysed sample of nickel-affinity purified PTA was injected onto the column at a flow rate of 0.3ml/min. Eluted protein was monitored at 280nm, and fractions were assayed for enzyme activity.

4.2.2.2 Ion-exchange chromatography

Ion-exchange chromatography was used to purify the native PTA from *G. thermoglucosidasius* TM242 cells. It was carried out using a 5ml HiTrap™ Q HP (anion) column on an ÄKTA explorer FPLC system. The pI values for the full-length PTA and the truncated PTA were predicted to be 5.49 and 5.66 respectively (ProtParam tool, ExPASy). The soluble protein fraction of *G. thermoglucosidasius* TM242 cell lysate prepared in 50mM

Epps, pH 7.0, was injected onto the column at a flow rate of 0.5ml/min. Bound protein was eluted by a gradual increase of a 50mM Epps buffer, pH 7.0, with 1M NaCl over 120min.

4.2.2.3 Dye-ligand affinity chromatography

Dye-ligand affinity chromatography uses immobilised ligands that resemble natural substrates to purify proteins. Reactive Red 120 is an analogue of $NADP^+$ and CoA, and an agarose-bound form was used to purify PTA from *G. thermoglucosidasius* TM242. The Reactive Red 120-agarose was loaded onto a chromatography column (Bio-Rad, Hertfordshire, UK) and equilibrated with 50mM Epps buffer, pH 7.0. The fractions collected from the anion exchange purification of PTA from *G. thermoglucosidasius* TM242 were pooled and loaded onto the column. Bound protein was eluted with 50mM Epps buffer, pH 7.0, containing 0.5mM CoA.

4.2.3 Enzyme characterisation

4.2.3.1 Dynamic light scattering

Dynamic light scattering (DLS) was used to determine the relative molecular mass of the recombinant PTA. The nickel-affinity purified sample (50µl) was loaded into a low volume quartz cuvette and heated to either 25°C or 60°C. The hydrodynamic diameter of the protein was measured using a Zetasiser Nano S from Malvern Instruments Ltd (Malvern, UK), and the relative molecular mass was calculated using the Protein utilities tool.

4.2.3.2 Dilution effects

To measure the catalytic activity of purified recombinant PTA by the spectrophotometric method described in Section 2.7.7.3, the sample required diluting 50- to 200-fold. The resulting activity loss was investigated by diluting samples in 50mM Tris-HCl, pH 8.0, with 150mM NaCl, at either 4°C or at room temperature; 50mM Hepes, pH 8.0, with 150mM NaCl, at room temperature; 50mM Tris-HCl, pH 8.0, and 150mM NaCl, with the addition of either 3mM 2-mercaptoethanol or 10mg BSA/ml. Immediately after dilution the sample was assayed for activity with 50mM phosphate and 0.14mM acetyl-CoA as described in Section 2.7.7.3. The diluted sample was either kept on ice or at room temperature as described, for three hours, and then assayed as before to determine the percentage of activity remaining.

4.2.3.3 Kinetic properties

Kinetic properties were determined for full-length and truncated PTA by the method described in Section 3.3.1.2.

4.2.3.4 Optimum temperature assays

The optimum temperatures of the recombinant PTA enzymes were determined by using a modified version of the assay described in Section 2.7.7.3. The assays were carried out in a 1ml quartz cuvette at temperatures between 30°C and 80°C. The buffer, 50mM sodium phosphate, pH 7.0, was pre-warmed to the desired temperature. The temperature of the reaction mixture was measured immediately after the rate was recorded.

4.2.3.5 Thermostability assays

Thermal inactivation assays were conducted in a Mastercycler PCR machine (Eppendorf, Hertfordshire, UK) using thin-walled 0.5ml Eppendorf tubes. The sample was heated in the PCR machine at various temperatures, with the lid set to 105°C to prevent evaporation. Sample tubes were removed periodically and cooled on ice. The sample was diluted in 50mM Tris-HCl, pH 8.0, 150mM NaCl, containing 5mg BSA/ml before the activity was measured as described in Section 2.7.7.3, with 50mM phosphate and 0.14mM acetyl-CoA.

4.2.3.6 Substrate specificity

The substrate specificity of the truncated PTA enzyme was determined by measuring the activity with 50mM phosphate and 0.1mM acyl-CoA. The acyl-CoAs tested were acetyl-CoA, propionyl-CoA and butyryl-CoA prepared using the method described in Section 2.7.7.1. The activity with acetyl-CoA was normalised to 100%.

4.2.4 Mass spectrometric analysis

4.2.4.1 Mass spectrometric analysis

The mass spectrometry (MS) experiments in Section 4.3.3 were performed and analysed by Dr Anneka Lubben (Chemistry Department, University of Bath). A sample of nickel-affinity purified PTA was subjected to buffer exchange by using spin filters. This exchanged the buffer for Milli-Q water, and concentrated the protein sample to 13mg/ml. Acetic acid was added to a final concentration of 0.1% (v/v) before samples were loaded onto the

MicrOTOF™ mass spectrometer coupled with an electrospray source (ESI-TOF) (Bruker Daltonics, Bremen, Germany). Data were analysed using the Bruker Compass DataAnalysis 4.0 software.

The MS experiments in Section 4.3.6 were performed and analysed by Dr Catherine Botting at the BSRC Mass Spectrometry and Proteomics Facility, St Andrews University (Botting C.H. 2013). The in-gel tryptic digest MS was performed by cutting each gel band into 1mm cubes. These were then subjected to in-gel digestion using a ProGest Investigator in-gel digestion robot (Genomic Solutions, Ann Arbor, USA) using standard protocols. The gel cubes were destained by washing with acetonitrile and subjected to reduction and alkylation before digestion with trypsin at 37°C. The peptides were then extracted with 10% (v/v) formic acid. The digest solution (0.5ml) was applied to the MALDI target along with alphacyano-4-hydroxycinnamic acid matrix (0.5ml, 10mg/ml in 50:50 acetonitrile:0.1% TFA) and allowed to dry. MALDI MS was acquired using a 4800 MALDI TOF/TOF Analyser (Applied Biosystems, Foster City, USA) equipped with an Nd:YAG 355nm laser and calibrated using a mixture of peptides. The most intense peptides (up to 15) were selected for MS-MS analysis and the combined MS and MS-MS data were analysed using GPS Explorer (Applied Biosystems) to interface with the Mascot 2.1 search engine (Matrix Science, London, UK), against the UniProt (Swiss-Prot and TREMBL combined) database. No species restriction was applied. The data were searched with tolerances of 100ppm for the precursor ions and 0.5Da for the fragment ions, trypsin as the cleavage enzyme, assuming up to one missed cleavage, carbamidomethyl modification of cysteines as a fixed modification and methionine oxidation selected as a variable modification.

The intact-MS experiments in Section 4.3.6 were also performed by Dr Botting. The protein sample was desalted on-line through a ChromXP C4-CL trap and column (Eskigent), using a nanoLC Ultra 2D plus loading pump and nanoLC AS-2 autosampler equipped with a nanoflex cHiPLC chip based chromatography system (Eskigent). The protein was eluted with a gradient of increasing acetonitrile, containing 0.1% formic acid (5-40% acetonitrile in 5min, 40-95% in a further 1min, followed by 95% acetonitrile to clean the column, before re-equilibration to 5% acetonitrile. The eluent was sprayed in to a TripleTOF 5600 electrospray tandem mass spectrometer (ABSciex). An envelope of multiply charged signals was obtained and deconvoluted using MaxEnt1 software to give the molecular mass of the protein.

4.3 Results

4.3.1 Cloning

The phosphotransacetylase (*pta*) gene annotated by ERGO™ was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA (Figure 4.1). The 1,062bp DNA fragment was excised from the gel, cloned into the pGEM®-T vector and transformed into *E. coli* JM109 electrocompetent cells as described in Section 4.2.1. The gene was sub-cloned into the pET28a(+) expression vector using restriction enzymes *Nde*I and *Xho*I. Plasmid DNA was purified, and sequencing confirmed that the correct nucleotide sequence was present.

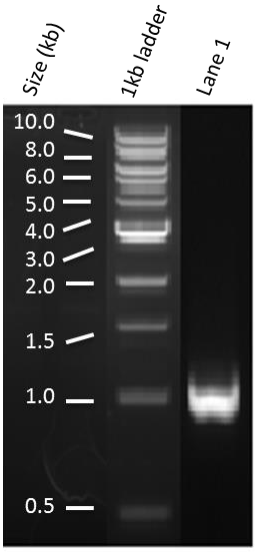


Figure 4.1 Agarose gel electrophoresis of the *pta* amplicon. The 1,062bp PCR-amplified gene (lane 1) was visualised alongside the 1kb ladder (NEB, Massachusetts, US).

4.3.2 Recombinant expression and purification

The putative *pta* gene was expressed in *E. coli* BL21(DE3) cells using the inducible T7 expression system from pET28a(+). This should yield a PTA protein with a poly(His)-tag at the N-terminus with a predicted molecular mass of 40.8kDa (Figure 4.7). His-tagged PTA was purified from a soluble cell lysate by nickel-affinity chromatography. SDS-PAGE revealed two major protein species co-purified by this method, one at approximately 43kDa, and a second around 34-35kDa (Figure 4.2).

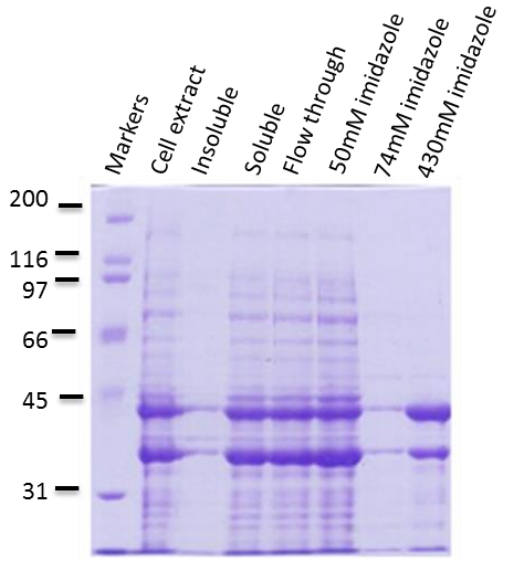


Figure 4.2 SDS-PAGE analysis of recombinant phosphotransacetylase (PTA). His-tagged PTA with an expected molecular mass of 40.8kDa was purified from *E. coli* BL21(DE3) cell lysate by nickel-affinity chromatography. The soluble fraction of cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and PTA was purified by washes with increasing imidazole concentrations. The size of the marker proteins are in kDa.

Size-exclusion chromatography of a partially purified sample was used in an attempt to identify the two protein species (Figure 4.3). However, this did not separate the two proteins. In addition, the PTA activity was detected in the void volume, suggesting PTA was forming large aggregated bodies of greater than 1MDa (Figure 4.4). The sample was also analysed by dynamic light scattering (DLS) at 25°C and 60°C (data not shown). However, these data also suggested the sample contained large aggregated bodies of between 1 and 7 MDa.

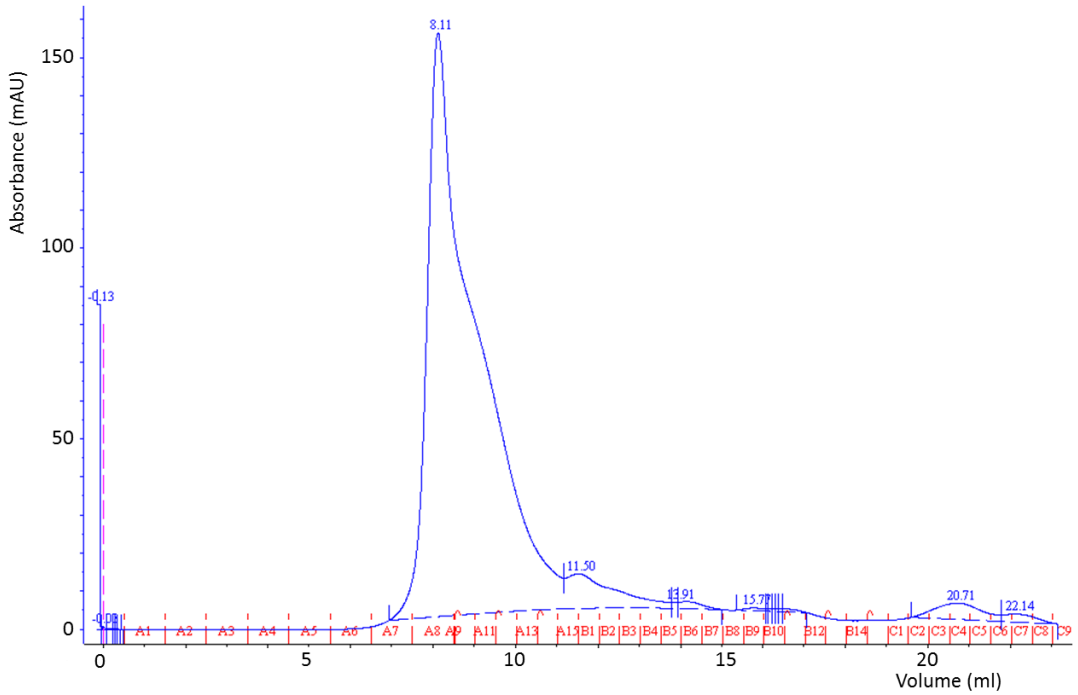


Figure 4.3 Size-exclusion chromatogram for recombinant PTA. The dark blue line is the absorbance detected at 280nm (mAU).

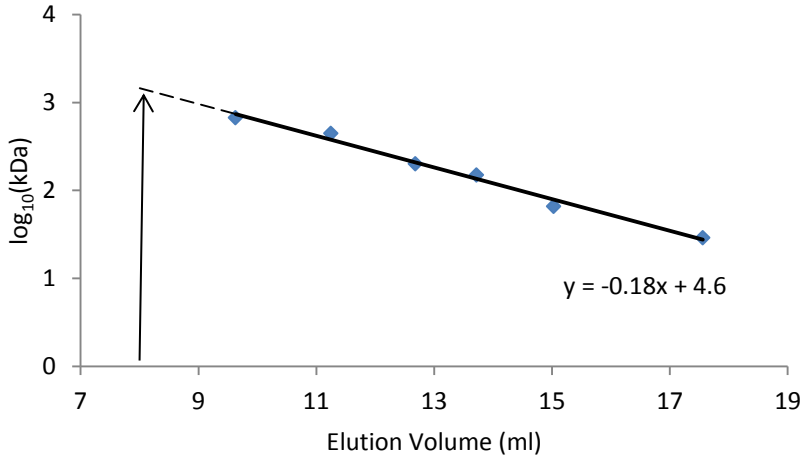


Figure 4.4 The relationship between elution volume on Superdex™ 200 10/300 GL gel filtration column and log(molecular weight). The gel filtration standards range from 29 to 669 kDa. The molecular weight of recombinant PTA was estimated by extrapolating back to the elution volume of PTA (arrow).

4.3.3 Mass spectrometry

Mass spectrometry (MS) determined the two protein species to be 40.68kDa and 35.06kDa in size (Figure 4.5). The larger of the two proteins correlates with the predicted molecular weight of the His-tagged PTA lacking the N-terminal Met (40.68kDa), as predicted by ExPASy's ProtParam tool (Gasteiger *et al.* 2005).

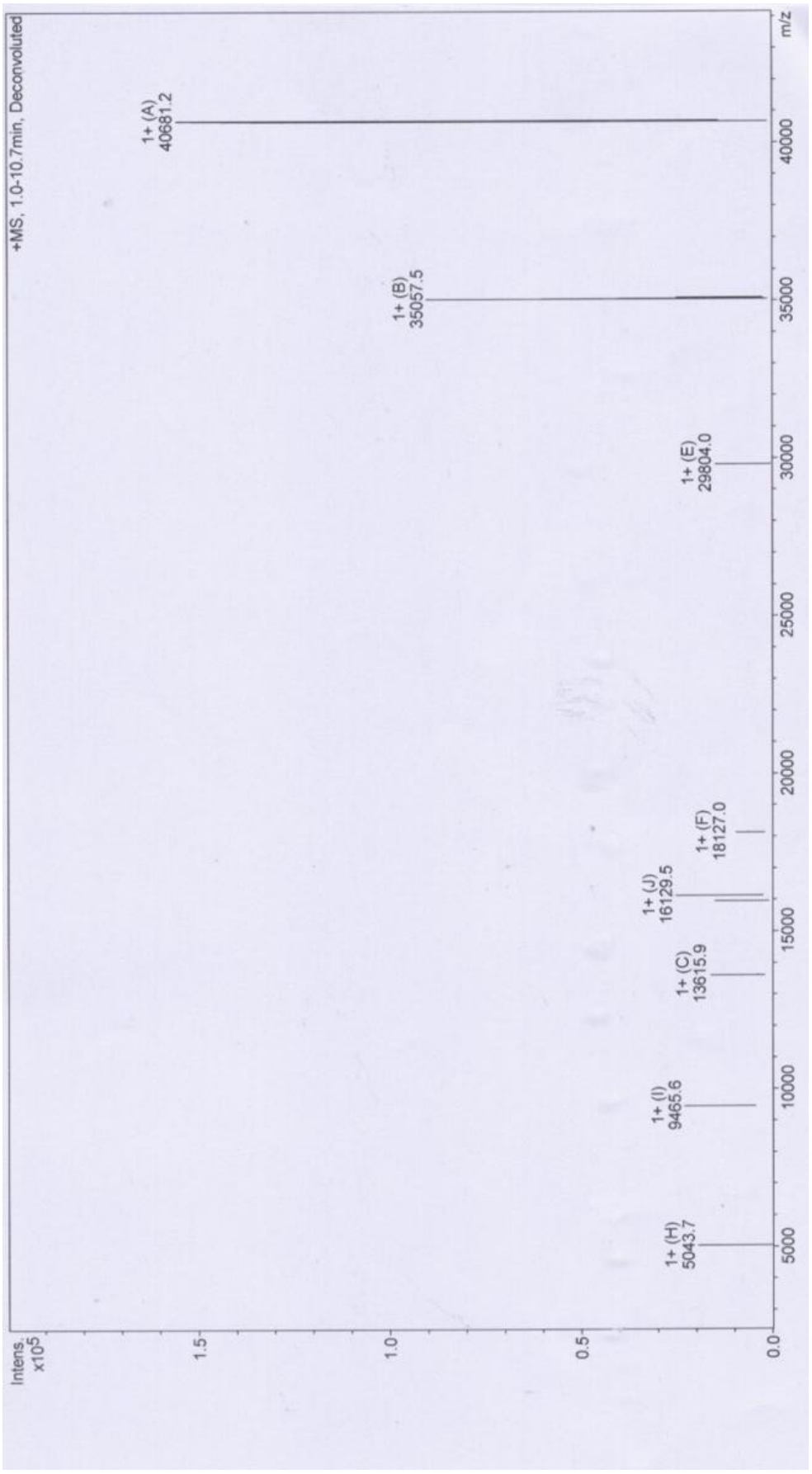


Figure 4.5 Mass spectrum of the nickel-affinity purified phosphotransacetylase (PTA).

Sequence alignment of PTA from *G. thermoglucosidasius* TM242 suggests the protein is greater than 70% identical and 80% similar to PTA enzyme's from the closely related *G. thermoglucosidasius* C56-YS93 and *Bacillus subtilis* 168 strains. However, the TM242 PTA has an extra 29 to 30 residues at its N-terminus (Figure 4.6).

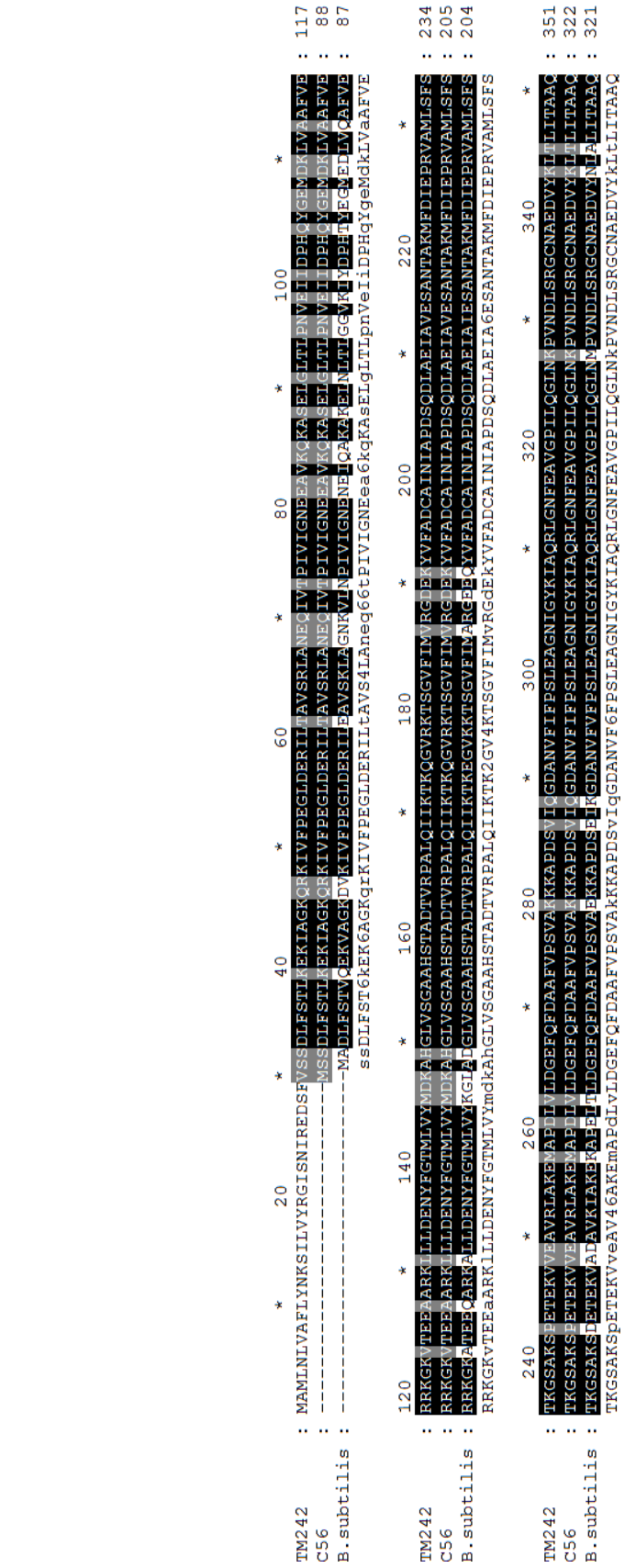


Figure 4.6 Sequence alignment of phosphotransacetylase (PTA) enzymes from *G. thermoglucosidasius* TM242, *G. thermoglucosidasius* C56-YS93 (UniProtKB accession number: F8CVL9) and *Bacillus subtilis* 168 (UniProtKB accession number: P39646). Dark shaded sequences indicate that all three residues at that position are identical while light shaded sequences indicate that two of the three residues at that position are identical.

The molecular weight of *G. thermoglucosidasius* TM242 His-tagged PTA without the first 29 and 30 residues was predicted to be 35.16kDa and 35.06kDa, respectively. The 30th residue in the His-tagged PTA sequence is a Val (Figure 4.6), encoded by GTG. GTG codons are the second most common translation initiators in *E. coli* (Blattner *et al.* 1997), and are translated as Met in such a case. It was hypothesised that a truncated version of the PTA enzyme could be synthesised in *E. coli* from this site. The N-terminal Met could be subsequently cleaved, yielding a 35.06kDa protein that could co-purify on a nickel affinity chromatography column due to the aggregating nature of PTA, despite the lack of an N-terminal His-tag. The presence of an *E. coli* ribosome binding site 5'- of the GTG codon supports this hypothesis.

MGSSHHHHHHSSGLVPRGSHMMAMLNLVAFLYNKSILVYRGISNIREDSFVSSDLFSTLK
EKIAGKQRKIVFPEGLDERILTAVSRLANEQIVTPIVIGNEEAVKQKASELGLTLPNVEI
IDPHQYGEMDKLVAAFVERRKGKVTEEAARKLLLDENYFGTMLVYMDKAHGLVSGAAHST
ADTVRPALQIIKTKQGVRKTSGVFIMVRGDEKYVFADCAINIAPDSQDLAEIAVESANTA
KMFDIEPRVAMLSFSTKGSAKSPETEKVVEAVRLAKEMAPDLVLDGEFQFDAAFVPSVAK
KKAPDSVIQGDANVFIFPSLEAGNIGYKIAQRLGNFEAVGPILQGLNKPVNDLSRGCNAE
DVYKLTLITAAQSL

Figure 4.7 The amino acid sequence of the full-length PTA expressed from the pET-28a(+) construct. The residues highlighted in grey are translated from the pET-28a(+) vector backbone sequence, and include the poly(His)-tag. The Val and Ser highlighted in yellow is from where translation of the hypothesised truncated PTA is initiated.

4.3.4 Cloning of full-length and truncated phosphotransacetylase species

To test whether translation from the alternative GTG site was the cause of the second protein species, constructs were made in order to express the two forms of PTA identified: full-length PTA and truncated PTA. To create a construct to express the full-length PTA alone, the GTG initiator identified was mutated to an alternative Val codon, GTT, by site-directed mutagenesis. This construct should allow for the expression of His-tagged full-length PTA, but prevent initiation of translation from the GTG. SDS-PAGE revealed the presence of a single dominant band of approximately 44kDa (Figure 4.8) from the nickel-affinity purification of full-length PTA. This purified sample had PTA-specific activity.

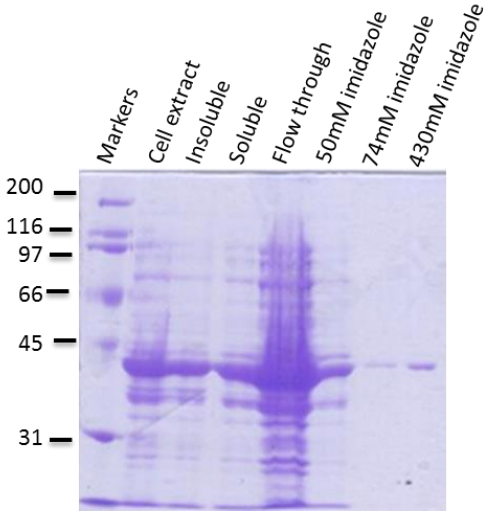


Figure 4.8 SDS-PAGE analysis of recombinant full-length PTA. His-tagged full-length PTA with an expected molecular mass of 40.8kDa was purified from *E. coli* BL21(DE3) cell lysate by nickel-affinity chromatography. The soluble fraction of cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and full-length PTA was purified by washes with increasing imidazole concentrations. The sizes of the marker proteins are in kDa.

To create a construct that would express the truncated PTA, the gene encoding the truncated protein was PCR-amplified and ligated into pET-28(a)+ using the restriction enzymes *Nde*I and *Xho*I. The GTG encoding the Val was mutated to ATG, to encode for a Met. His-tagged truncated PTA (Figure 4.9) was expressed in *E. coli* BL21(DE3) cells using the T7 expression system. The truncated PTA protein was purified by nickel-affinity chromatography, and SDS-PAGE revealed a single band of around 38kDa (Figure 4.10). This purified sample had PTA-specific activity, suggesting the second species initially observed (Figure 4.2) is a result of the synthesis of a truncated PTA.

MGSSHHHHHHSSGLVPRGSHMMSSDLFSTLKEKIAGKQRKIVFPEGLDERILTAVSRLAN
EQIVTPIVIGNEEAVKQKASELGLTLPNVEIIDPHQYGEMDKLVAAFVERRKGKVTEEAA
RKLLLDENYFGTMLVYMDKAHGLVSGAAHSTADTVRPALQIIKTKQGVRKTSGVFIMVRG
DEKYVFADCAINIAPDSQDLAEIAVESANTAKMFDIEPRVAMLSFSTKGSAKSPETEKVV
EAVRLAKEMAPDLVLDGEFQFDAAFVPSVAKKKAPDSVIQGDANVFIFPSLEAGNIGYKI
AQRLGNFEAVGPILQGLNKPVNDLSRGCNAEDVYKLTLITAAQSL

Figure 4.9 The amino acid sequence of truncated PTA expressed from the pET-28a(+) construct. The residues highlighted in grey are translated from the pET-28a(+)vector backbone sequence, and include the poly(His)-tag. The Met highlighted in yellow is where the hypothesised truncated PTA is translated from after the GTG in the native DNA sequence has been changed to ATG in this construct.

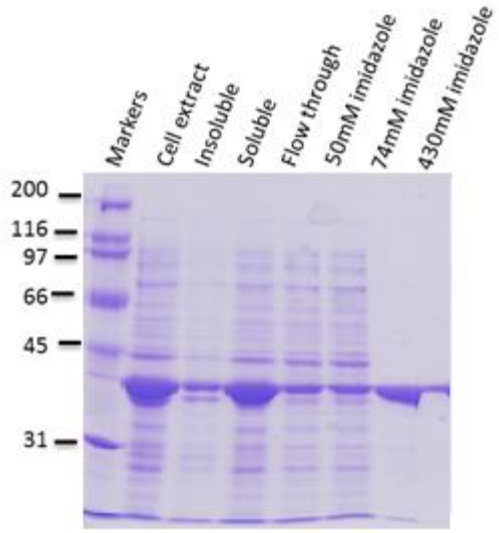


Figure 4.10 SDS-PAGE analysis of recombinant truncated PTA. His-tagged truncated PTA with an expected molecular mass of 37.5kDa was purified from *E. coli* BL21(DE3) cell lysate by nickel-affinity chromatography. The soluble fraction of the cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and truncated PTA was purified by washes with increasing imidazole concentrations. The sizes of the marker proteins are in kDa.

4.3.5 Characterisation of the two forms of recombinant phosphotransacetylase

4.3.5.1 Dilution effects

Both the full-length PTA and the truncated PTA purified by nickel-affinity chromatography required diluting between 50- and 200-fold in order to measure initial rates spectrophotometrically. Upon dilution of a sample into a 50mM Tris buffer, 40% of enzymic activity was lost after a 3h incubation period (Tris, 4°C; Figure 4.11). Neither diluting the sample in a 50mM Hepes buffer nor incubation at room temperature prevented activity loss. The addition of 3mM 2-mercaptoethanol, which can prevent the oxidation of thiol groups, was unable to prevent activity loss. However, the addition of bovine serum albumin (BSA), at 10mg BSA/ml, to the 50mM Tris buffer led to less than 3% activity loss after 3h incubation period. The observation that undiluted purified PTA activity was stable over the same incubation period suggests that the reduction in activity was due to a loss of protein stability in diluted samples. Further experiments established that addition of 5mg BSA/ml to the dilution buffer was sufficient to maintain greater than 97% activity for 3h post-dilution. In addition, 90% of activity was maintained 24h post-dilution with 5mg BSA/ml, compared to less than 35% activity without the addition of BSA.

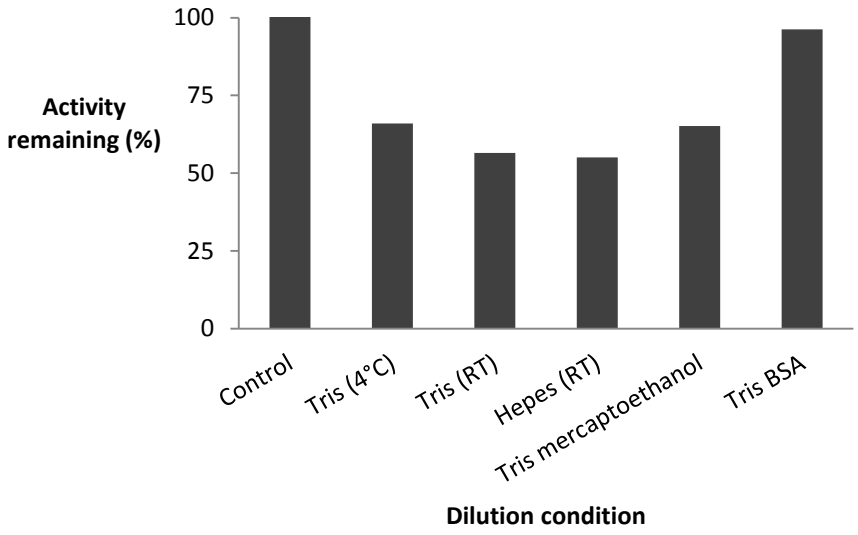


Figure 4.11 Relative phosphotransacetylase activity remaining 3h post sample dilution. Purified recombinant PTA activity was measured immediately after dilution, and the remaining activity is shown after a 3h incubation period. RT indicates the sample was incubated at room temperature; all other samples were incubated at 4°C. 3mM 2-mercaptoethanol and 10mg BSA/ml were added as indicated. The control sample was diluted at 0h and 3h, then assayed to test whether activity loss occurs without dilution.

4.3.5.2 Kinetic properties

To compare the kinetic properties of the two recombinant PTA species, a matrices approach was undertaken as described in Section 3.3.1.2. The purified proteins were diluted in a 50mM Tris buffer, pH 8.0, containing 5mg BSA/ml, and the ionic strength of the assay was maintained at 0.1M with the addition of NaCl. The data were analysed by the non-linear fit method from the enzyme kinetics module of SigmaPlot for both the full-length PTA (Figure 4.12 and Figure 4.13) and the truncated PTA (Figure 4.14 and Figure 4.15). The kinetic parameters are summarised in Table 4.2.

Table 4.2 Kinetic parameters determined for full-length and truncated PTAs. Values are given, plus or minus standard errors.

Enzyme	Substrate	V_{max} (U/mg)	K_m (mM)
Full-length PTA	acetyl-CoA	1,360 (± 90)	0.070 (± 0.013)
	phosphate	1,350 (± 130)	4.2 (± 1.5)
Truncated PTA	acetyl-CoA	1,580 (± 140)	0.059 (± 0.017)
	phosphate	1,580 (± 100)	3.5 (± 0.9)

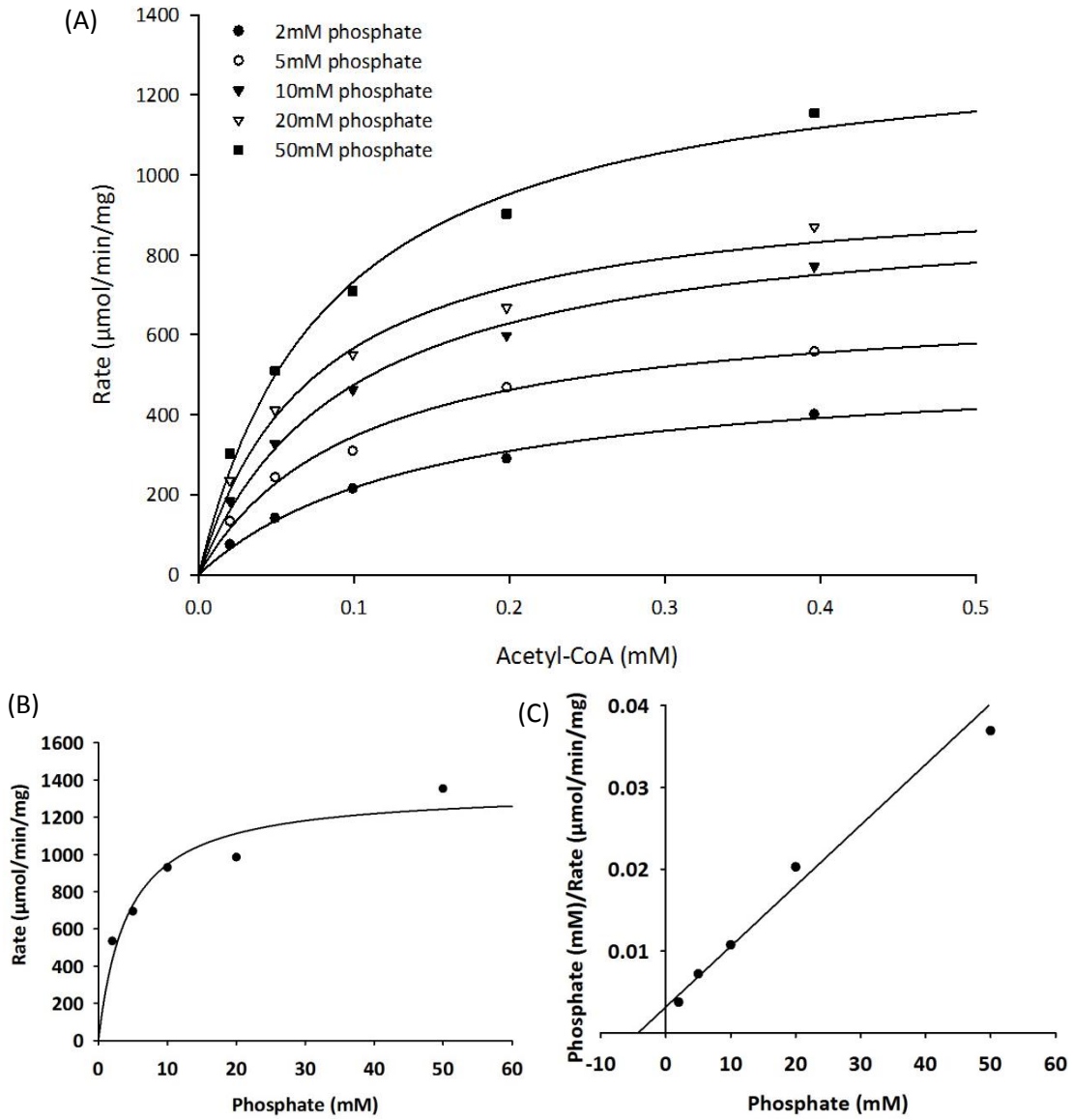


Figure 4.12 Full-length PTA activity. The relationship between specific activity (U/mg) and acetyl-CoA concentration at a variety of fixed phosphate concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true K_m for phosphate at saturating acetyl-CoA concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

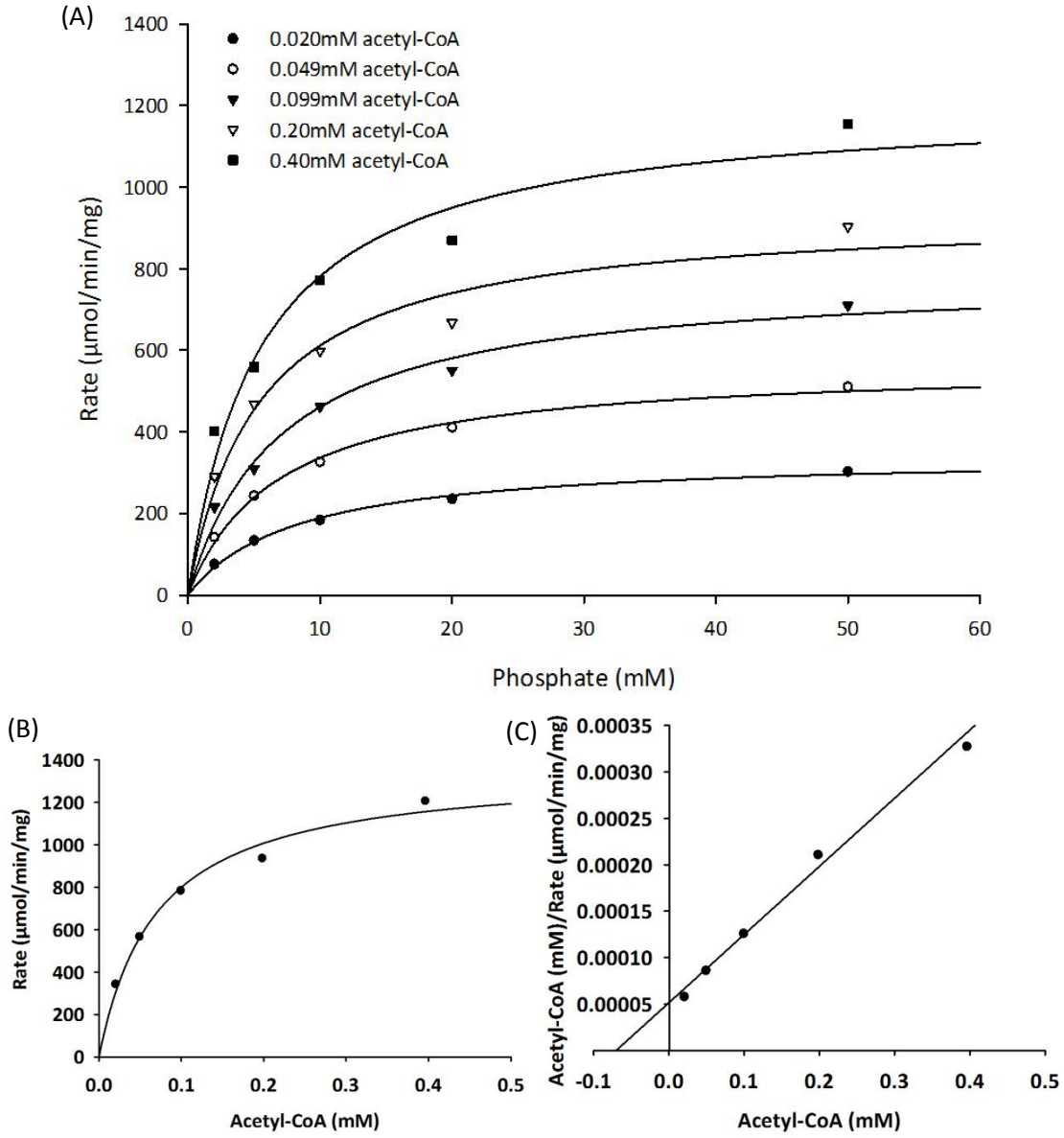


Figure 4.13 Full-length PTA activity. The relationship between specific activity (U/mg) and phosphate concentration at a variety of fixed acetyl-CoA concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true $\boldsymbol{K_m}$ for acetyl-CoA at saturating phosphate concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

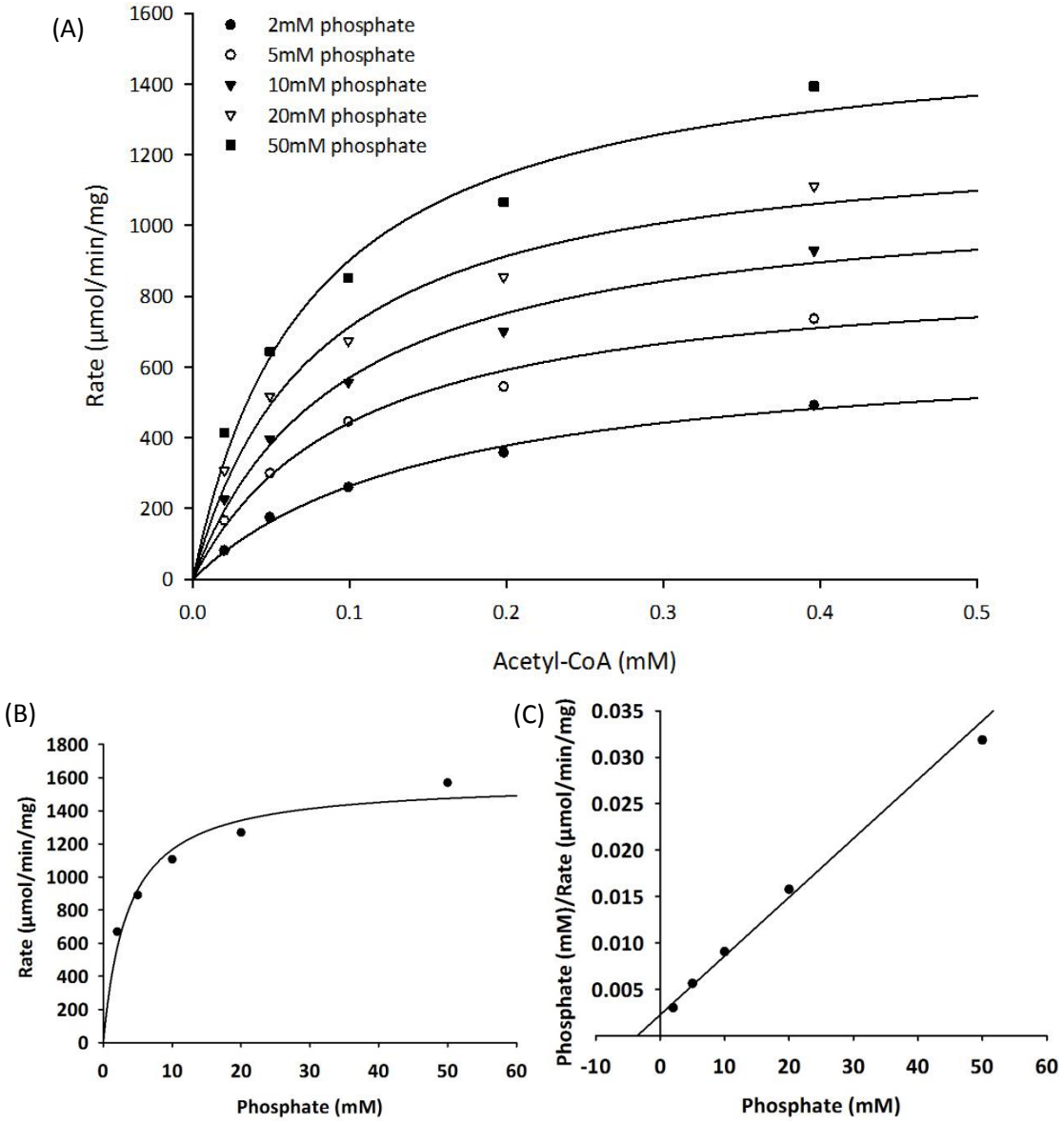


Figure 4.14 Truncated PTA activity. The relationship between specific activity (U/mg) and acetyl-CoA concentration at a variety of fixed phosphate concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true $\boldsymbol{K_m}$ for phosphate at saturating acetyl-CoA concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

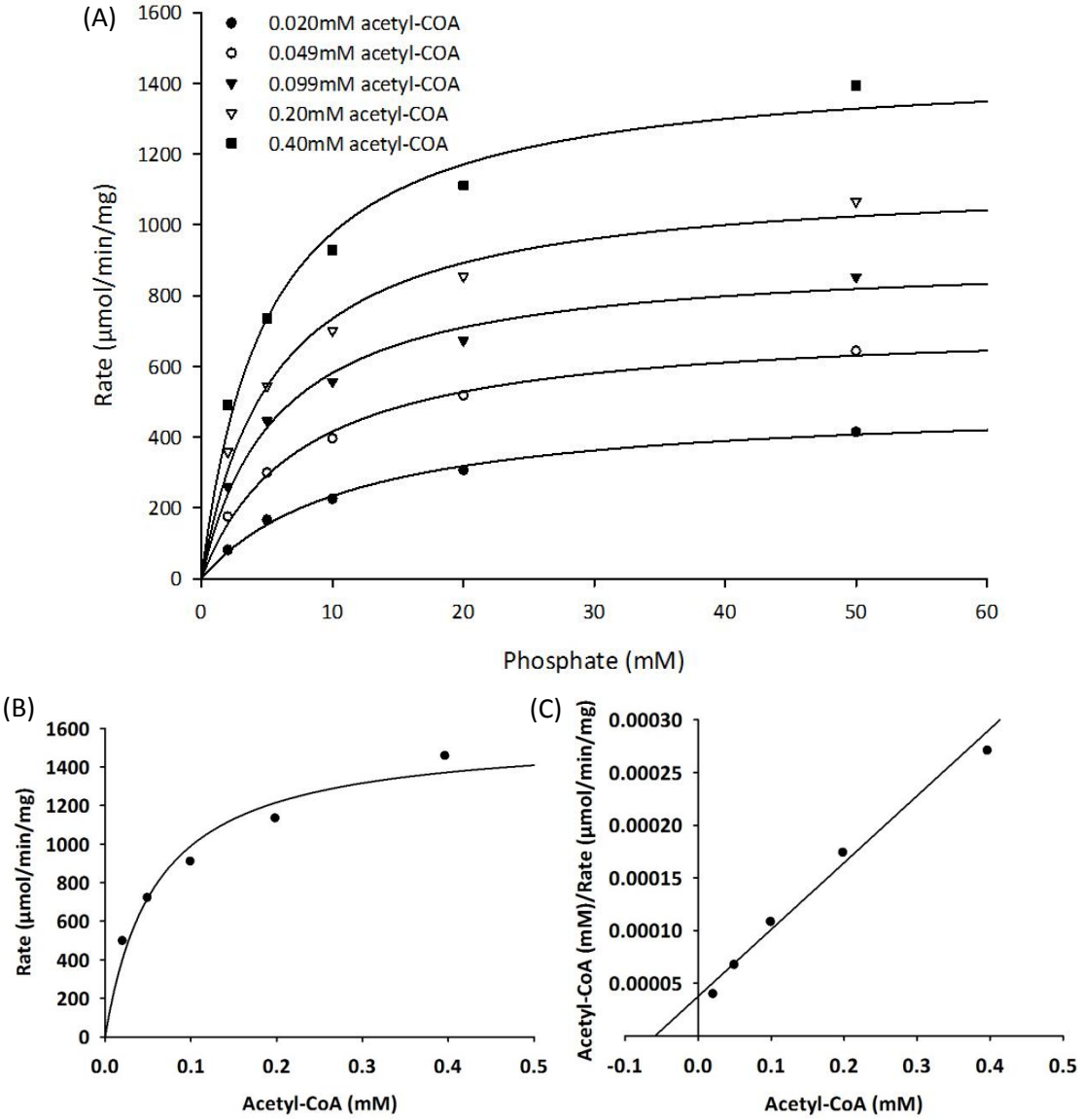


Figure 4.15 Truncated PTA activity. The relationship between specific activity (U/mg) and phosphate concentration at a variety of fixed acetyl-CoA concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true K_m for acetyl-CoA at saturating phosphate concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

4.3.5.3 Temperature optimum

To characterise the two forms of recombinant PTA further, their properties in relation to temperature were investigated. The temperature optimum for both forms was between 60°C and 62°C, and the temperature profiles overlapped (Figure 4.16).

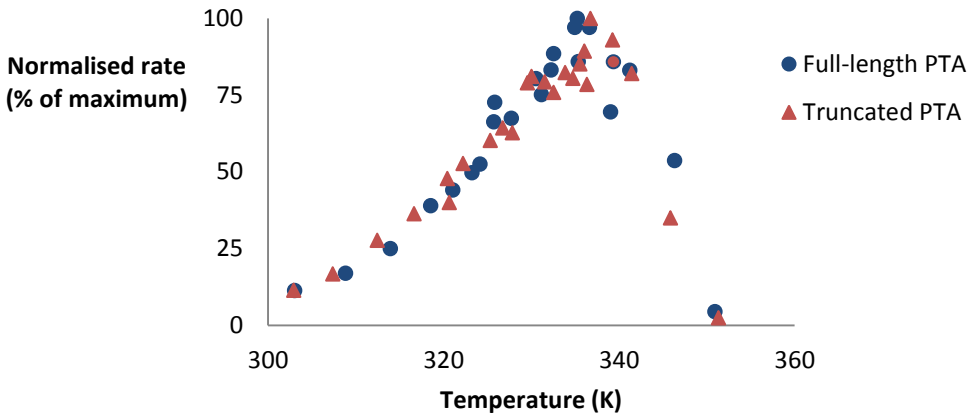


Figure 4.16 The relationship of temperature (K) and enzyme activity for full-length PTA and truncated PTA.

4.3.5.4 Thermostability

The thermostability of full-length PTA and truncated PTA was determined by incubating undiluted purified aliquots at a variety of temperatures. These aliquots were then diluted into a 50mM Tris buffer containing 5mg BSA/ml, and the PTA activity was measured. The data suggest that the truncated form is more thermostable than the full-length enzyme (Figure 4.17). The half-life of the truncated PTA at 65°C is 120min compared to 30min for the full-length PTA.

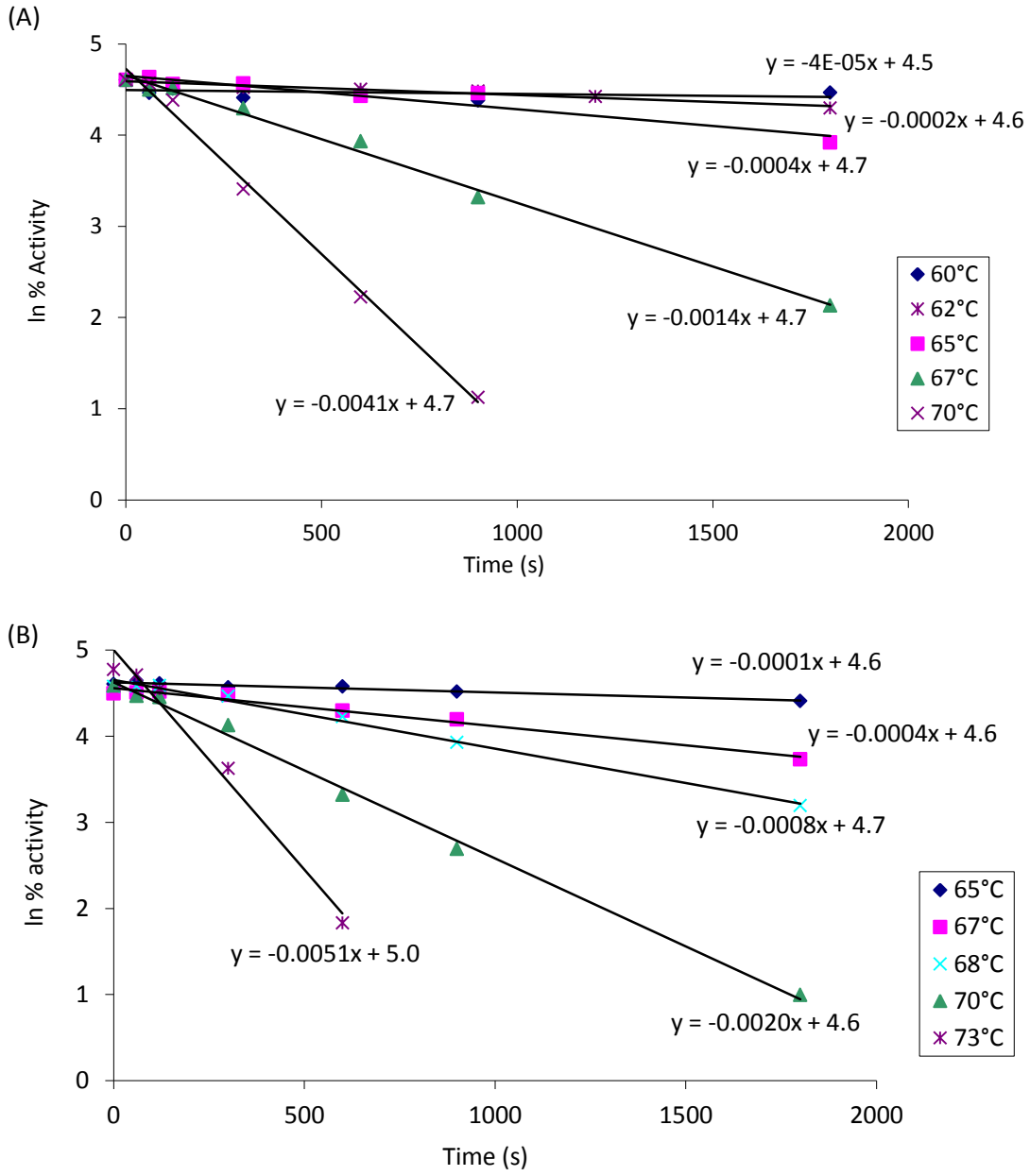


Figure 4.17 Comparison of enzyme thermostability. The relationship between ln(activity) and incubation time at a variety of temperatures for full-length PTA (A), and truncated PTA (B).

The rate constants for the thermal inactivations determined from Figure 4.17 can be used to calculate the activation energy of inactivation for each PTA using the Arrhenius equation:

$$k = Ae^{-Ea/RT}$$

where k is the rate constant, A is the pre-exponential factor, Ea is the activation energy, R is the Gas constant and T is the absolute temperature. The activation energy for the full-length PTA and the truncated PTA were determined to be 420kJ/mol and 470kJ/mol respectively (Figure 4.18).

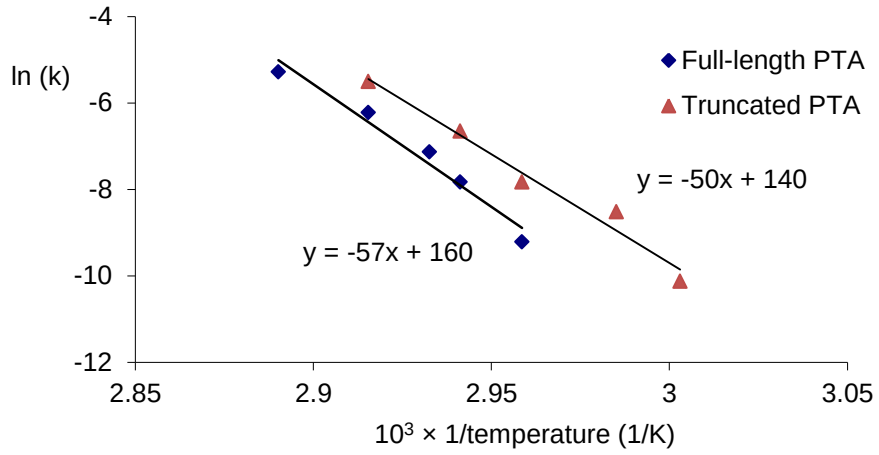


Figure 4.18 Arrhenius plot comparing the thermal inactivation of full-length PTA and truncated PTA.

4.3.6 Which protein species is native to *G. thermoglucosidasius*?

To determine the form(s) of PTA natively produced, the enzyme was purified from glucose grown *G. thermoglucosidasius* TM242 cells. Anion exchange chromatography was used for the first step. A TM242 cell lysate prepared in 50mM Epps, pH 7.0, with the addition of protease inhibitors was loaded onto a HiTrap™ 5ml Q HP anion exchange column. Once the initial flow-through had been washed through bound protein was eluted by increasing the NaCl concentration in the buffer. The elution volume of native PTA was determined by assaying the 2ml fractions collected (Figure 4.19).

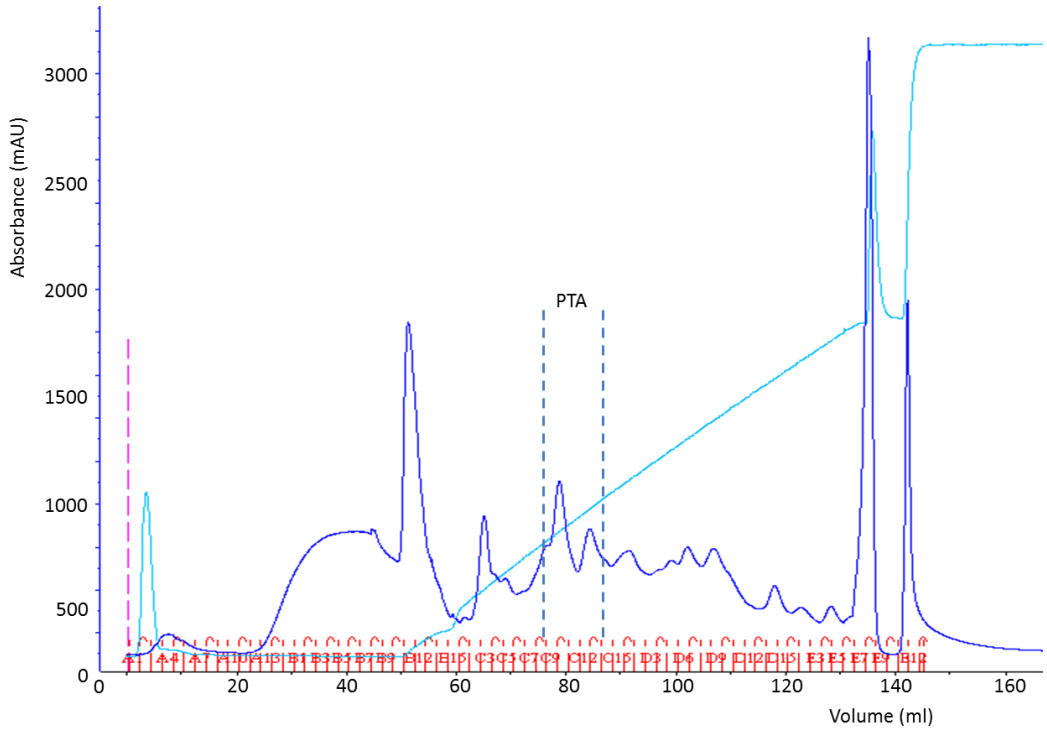


Figure 4.19 Anion exchange chromatograph for the purification of PTA from *G. thermoglucosidasius* TM242. The dark blue line is the absorbance at 280nm (mAU), the light blue line represents the increase in NaCl concentration, and the fractions where PTA activity was detected are indicated with the dashed lines.

The fractions containing PTA activity were pooled and the PTA was further purified by using dye-affinity chromatography on Reactive Red 120-agarose. After this two-step purification around ten different protein species were observable by SDS-PAGE (Figure 4.20). The purity of the native PTA had increased 75-fold, and 74% of the PTA activity had been recovered. Two of the proteins purified from the *G. thermoglucosidasius* TM242 cell lysate migrate to the same point as purified recombinant PTA (which, as identified in Section 4.3.2, includes the His-tagged full-length PTA and the non-His-tagged truncated PTA).

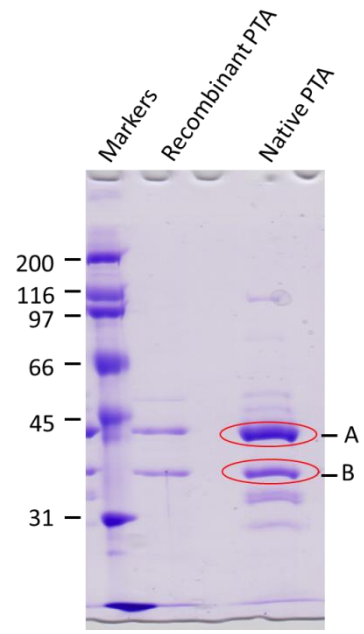


Figure 4.20 SDS-PAGE of recombinant PTA purified from *E. coli* BL21(DE3) cells and native PTA partially purified from *G. thermoglucosidasius* TM242. The two dominant bands in the recombinant PTA sample are the full-length His-tagged PTA and the non-His-tagged truncated forms. Bands circled were excised and analysed by tryptic mass spectrometry.

The native PTA sample was analysed by mass spectrometry (MS), both by intact-MS and by in-gel tryptic digest MS. The intact-MS revealed three abundant protein species of sizes 30.8kDa, 35.1kDa and 42.3kDa (Figure 4.21). No protein was evident in the partially purified sample with a mass equivalent to that for the full-length PTA (approximately 38kDa). However, a protein with a mass equivalent to that of the truncated version was identified (35.06kDa).

Bands labelled A and B (Figure 4.20) were excised from the gel and analysed by in-gel tryptic MS. The peptide fragments identified by this method for band A showed similarity to phosphoglycerate kinase (Mr = 42,342) and not to the PTA. The fragments identified for the protein in band B showed similarity for PTA, and a mass corresponding to the one fragment that is unique for the truncated PTA was identified (see Appendix Figure 10.2 for unique peptide sequence). No peptides were identified with a mass corresponding to any of the four unique fragments for the full-length PTA (see Appendix Figure 10.1 for unique peptide sequences).

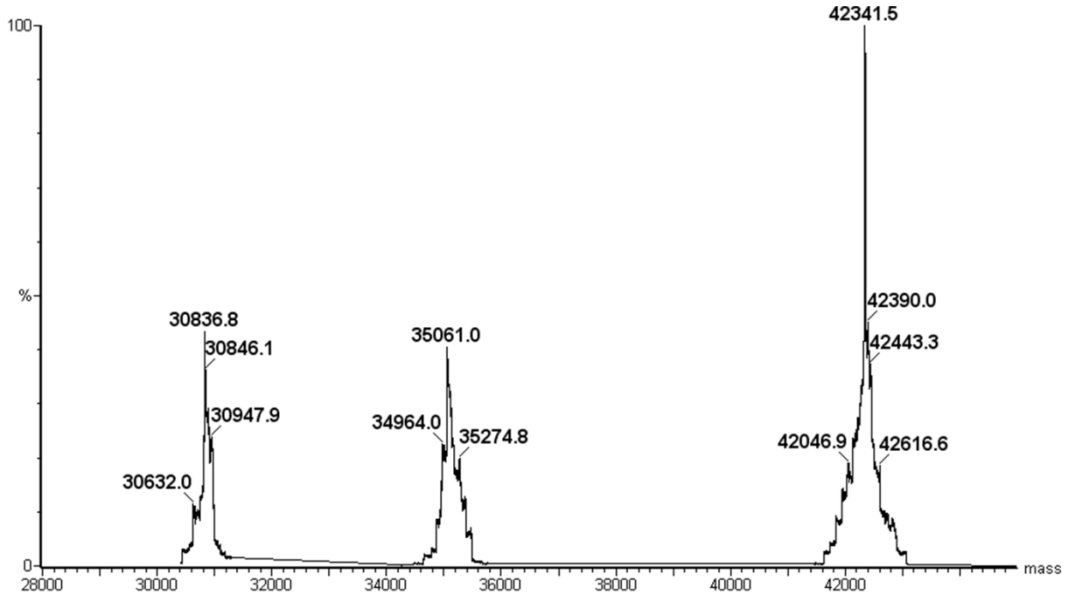


Figure 4.21 Intact mass spectrometry of the partially purified PTA from *G. thermoglucosidasius* TM242 cell lysate.

4.3.7 Substrate specificity

The substrate specificity of recombinantly expressed truncated PTA was determined from a nickel-affinity purified sample. Truncated PTA can catalyse the acyl-group transfer to orthophosphate with acetyl-CoA or propionyl-CoA (Figure 4.22). The acyl-group transfer to orthophosphate from butyryl-CoA was less than 0.1% of that measured with acetyl-CoA.

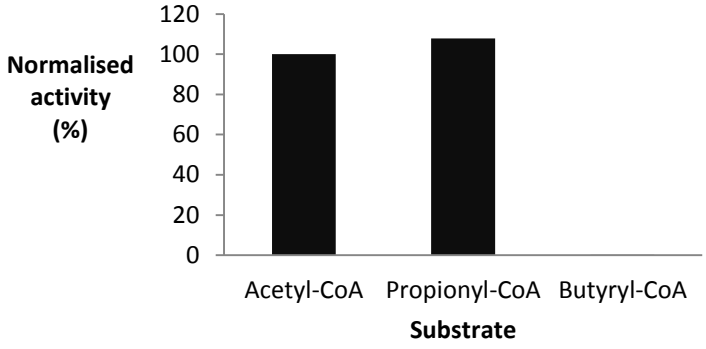


Figure 4.22 Substrate specificity of truncated PTA. Purified recombinant truncated PTA was assayed with 50mM phosphate and either 0.1mM acetyl-CoA, propionyl-CoA or butyryl-CoA. Activity with acetyl-CoA was normalised to 100%.

4.4 Discussion

Recent years have seen major advances in 'next-generation' sequencing technologies (Koboldt *et al.* 2013; MacLean *et al.* 2009), to an extent where an organism's

genome can be sequenced in a number of hours to days. However, the technologies and methods used to annotate the sequence, despite advancements, have not improved at the same pace. Genome annotation usually involves an automated process, followed by a manual curation (Stothard and Wishart 2006). The automated process relies heavily on sequence identity with previously annotated genes and genomes of closely related species, but this method has many limitations associated with it. One such is that any errors associated with the homologous sequence may be passed on to the new sequence. In addition, automated processes usually identify sequences relating to specific genes but often do not include accurate details, such as ribosome binding sites, and terminators and initiators to both transcription and translation (Richardson and Watson 2013). To be fully certain that an annotated gene encodes a functional protein experimental evidence is required.

The *G. thermoglucosidasius* TM242 *pta* gene identified by the ERGO™ Integrated Genomics software was highly homologous to those from *G. thermoglucosidasius* C56-YS93 and *Bacillus subtilis* 168. A combination of a bioinformatics approach and experimental evidence identified two possible translational initiators for *pta*: an ATG codon and a less common GTG codon. Constructs were designed and made in order to characterise the full-length and truncated PTA versions, corresponding to proteins translated from the ATG and GTG codons, respectively.

It was confirmed that both the full-length and the truncated forms of PTA had phosphotransacetylase activity. The K_{m} values for the recombinant proteins were similar to those determined in Chapter 3 for the native PTA from *G. thermoglucosidasius* TM242 cell lysate. This suggests that the recombinant expression system produced enzymes comparable to the natively expressed PTA, and was a viable system for characterising the purified PTA from *G. thermoglucosidasius* TM242. In addition, the kinetic parameters determined in this study closely resemble those reported for *B. subtilis* PTA; Rado and Hoch (1973) partially purified PTA from *B. subtilis* and reported a specific activity of 1371 µmol acetyl-CoA cleaved/min/mg of protein, and K_{m} values of 0.06mM and 10mM for acetyl-CoA and orthophosphate, respectively. Shin *et al.* (1999) have characterised a His-tagged recombinant PTA from *B. subtilis* and report a specific activity of 1150 µmol acetyl-phosphate cleaved/min/mg, although this was in the acetyl-CoA forming direction.

The minimal difference in kinetic parameters between the full-length and truncated PTA enzymes from *G. thermoglucosidasius* TM242 does not give any clues as to the form produced *in vivo* by this organism. Likewise, the similar thermal properties determined in this study do not distinguish between them. However, caution should be exercised here when comparing these two proteins because both recombinant forms of PTA have 21 additional residues (including the six His residues of the poly(His)-tag) on their N-termini. This additional peptide, a result of the T7 promoter system of pET-28a(+), could change the characteristics of these two enzymes. Although, it is worth noting that the structure of the *B. subtilis* PTA has been solved and has revealed that the N-terminus is located at the surface of the protein (Xu *et al*. 2005). The *G. thermoglucosidasius* TM242 PTA is highly homologous to that of the *B. subtilis* PTA, suggesting that the N-terminus may also be located at the surface of the protein and that the additional N-terminal peptide may not significantly affect structure or function. To determine whether this is the case thrombin cleavage of the N-terminal peptide could have been performed prior to characterisation of these two enzymes, or the C-terminus could have been tagged by the use of a different vector.

The partial purification of the native PTA from *G. thermoglucosidasius* TM242 determined that the truncated version was the form produced *in vivo*, under the conditions used in this study. This suggests that either the full-length PTA is cleaved by post-translational modification, or the more likely scenario that the GTG acts as the translational initiator. Despite a high recovery, 25% of PTA-specific activity was unaccounted for after the purification, and therefore it is possible that the full-length protein was lost. From this study it cannot be concluded whether under certain conditions *G. thermoglucosidasius* produces the full-length PTA.

In conclusion this Chapter has determined that the putative *pta* gene in *G. thermoglucosidasius* TM242 does encode a PTA enzyme, and has improved the annotation by investigating the translational start sites. The truncated protein, translated from a GTG initiator, was determined to be the natively produced PTA. The kinetic properties determined in this Chapter are comparable to those published for PTA enzymes from similar organisms, and the substrate specificity has also been studied.

Chapter 5
Characterisation of recombinant acetate kinase

5.1 Introduction

Acetogenesis is the formation and secretion of acetate, which often occurs in the presence of excess carbon source. The enzymes phosphotransacetylase (PTA) and acetate kinase (AK) catalyse the conversion of acetyl-CoA to acetate, with the concomitant generation of ATP. PTA catalyses the reversible transfer of the acetyl group between CoA and orthophosphate (Equation 5.1), while the phosphorylation of ADP by acetyl-phosphate is catalysed by AK (Equation 5.2).

$$CH_3COCoA + HPO_4^{2-} \rightleftarrows CH_3COOPO_3^{2-} + CoA \quad (5.1)$$

$$CH_3COOPO_3^{2-} + ADP^{3-} \rightleftarrows CH_3COO^{-} + ATP^{4-} \quad (5.2)$$

Acetogenesis is important for several key reasons. Firstly, it helps replenish CoA pools, which facilitates rapid growth by providing the CoA for α-ketoglutarate conversion to succinyl-CoA in the citric acid cycle (CAC) (El-Mansi 2004). Secondly, it provides a significant amount of ATP in the absence of a fully functioning CAC, during fermentative growth for example (Wolfe 2005). Lastly, there is surmounting evidence that the intermediate of the PTA-AK pathway, acetyl-phosphate, can act as a global signal (McCleary *et al.* 1993).

Acetyl-phosphate possesses a larger $\Delta G°$ of hydrolysis (-43.3 kJ/mol) than ATP (-30.5 kJ/mol in complex with Mg^{2+}), which provides the basis for its role in global regulation. It has been reported that acetyl-phosphate can donate its phosphoryl group to a subset of response regulators of the family of two-component signal transduction pathways (McCleary *et al.* 1993; Wanner 1993; Wolfe 2005). These signalling pathways have been shown to be involved with controlling osmoregulation, nitrogen assimilation, pilus construction, flagella biogenesis, capsule biosynthesis, bio-film development and pathogenicity (Wolfe 2005; 2010).

In addition to phosphorylation, acetyl-phosphate has been linked to cellular regulation by acetylating certain lysine residues of proteins (Verdin and Ott 2013; Weinert *et al.* 2013). Protein acetylation is primarily believed to occur through the donation of an acetyl group from acetyl-CoA (Hu *et al.* 2010; Wang *et al.* 2010). However, through disruptions made to

the PTA-AK pathway, Weinert *et al*. (2013) demonstrated that the acetylation levels in *E. coli* were directly related to the organism's ability to produce acetyl-phosphate.

The primary function of the PTA-AK pathway in *G. thermoglucosidasius* TM242 has not been studied. By studying the *in vivo* activities of PTA and AK (Chapter 3), and the kinetic properties of these enzymes, a better picture of steady-state substrates levels can be built, which may elucidate the true function of this pathway.

This Chapter describes the cloning of the gene encoding acetate kinase (AK), the recombinant expression of AK in an *E. coli* system, and the characterisation of the recombinant enzyme in terms of kinetic properties. The cloning work in this Chapter was completed by supervised undergraduate student Kirstie Cleary.

5.2 Methods

5.2.1 Cloning

Cloning of the acetate kinase (*ak*) gene used general molecular biology techniques described in Section 2.6. Vector-specific primers were used for DNA sequencing to determine that the correct sequence had been cloned. pET-45b(+) clones were transformed into *E. coli* BL21(DE3) cells for protein expression.

5.2.1.1 *ak*-pET-45b(+)

The *ak* gene was amplified by the polymerase chain reaction (PCR) from *G. thermoglucosidasius* TM242 genomic DNA, using primers *ak*F1 and *ak*R1 (Table 5.1). The PCR product was purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *ak* gene was excised from the pGEM®-T vector by restriction digest with *Kpn*I and *Xho*I, agarose gel purified and ligated into pET-45b(+) that had been linearized with *Kpn*I and *Xho*I. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the construct.

Table 5.1 PCR primers used for *ak* cloning. Restriction sites are underlined.

Name	Sequence (5'-3')
*ak*F1	GGTACCATGATGATTAAGGGAGTTGTTTG
*ak*R1	CCTCGAGGTTAAGCAATATTCGCCAACCG

5.2.2 Protein purification

Recombinant protein was purified by nickel-affinity chromatography as described in Section 2.7.4. Imidazole was removed from samples of purified recombinant protein by overnight dialysis at 4°C into a 50mM Tris-HCl buffer, pH 8.0, containing 150mM NaCl. Purified recombinant AK was unstable after diluting to a concentration necessary for enzymatic assays (see Section 5.3.3.1). Therefore, protein samples were diluted in buffers containing bovine serum albumin (5mg BSA/ml).

5.2.3 Enzyme characterisation

5.2.3.1 Dilution effects

To measure the catalytic activity of purified recombinant AK by the spectrophotometric method described in Section 2.7.7.4, the sample required diluting 200- to 1000-fold. The resulting activity loss was investigated by diluting samples in 50mM Tris-HCl, pH 8.0, with and without the addition of 10mg BSA/ml. Immediately after dilution, the sample was assayed for activity with 3.0mM acetyl-phosphate and 2.0mM ADP as described in Section 2.7.7.4. The diluted sample was kept on ice for three hours, and then assayed as before to determine the percentage of activity remaining.

5.2.3.2 Kinetic properties

Kinetic properties were determined by the method described in Section 2.7.7.4.

5.3 Results

5.3.1 Cloning

The *ak* gene was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA (Figure 5.1). The 1,221bp DNA fragment was excised from the gel, cloned into the pGEM®-T vector and transformed into *E. coli* JM109 electrocompetent cells as described in Section 4.2.1.

The gene was sub-cloned into the pET45b(+) expression vector using restriction enzymes *Kpn*I and *Xho*I. Plasmid DNA was purified, and sequencing confirmed that the correct nucleotide sequence was present.

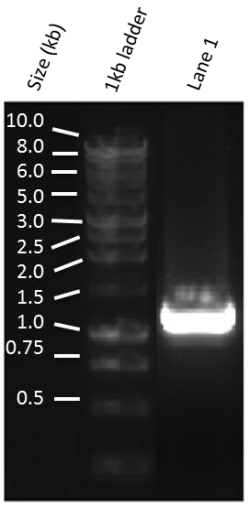


Figure 5.1 Agarose gel electrophoresis of the *ak* amplicon. The 1,221bp PCR-amplified gene (lane 1) was visualised alongside the 1kb ladder (NEB, Massachusetts, US).

5.3.2 Recombinant expression and purification

The putative *ak* gene was expressed in *E. coli* BL21(DE3) cells using the inducible T7 expression system from pET45b(+). This resulted in an AK protein with a poly(His)-tag at the N-terminus and a predicted molecular mass of 45.7kDa. His-tagged AK was purified from a soluble cell lysate by nickel-affinity chromatography. SDS-PAGE estimated the AK band to be 45kDa (Figure 5.2).

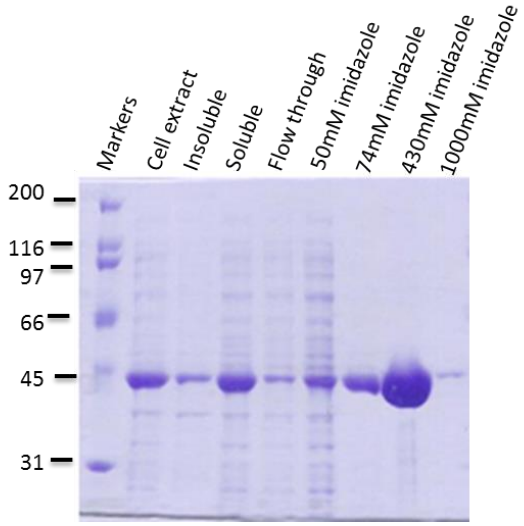


Figure 5.2 SDS-PAGE analysis of recombinant acetate kinase (AK). His-tagged AK with an expected molecular mass of 45.7kDa was purified from *E. coli* BL21(DE3) cell lysate by nickel-affinity chromatography. The soluble fraction of cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and AK was purified by washes with increasing imidazole concentrations. The size of the marker proteins are in kDa.

5.3.3 Characterisation of the recombinant acetate kinase

5.3.3.1 Dilution effects

Purified recombinant AK was subjected to buffer exchange by overnight dialysis to remove the imidazole, and to exchange the buffer to a 50mM Tris-HCl buffer, pH 8.0, containing 150mM NaCl. AK activity was measured by monitoring ATP production using the enzyme-linked assay system with hexokinase and glucose-6-phosphate dehydrogenase, as previously described in Section 2.7.7.4. The purified AK required significant dilution in order to measure initial rates spectrophotometrically. However, similar to recombinant PTA, 35% of enzymic activity was lost after a 3h incubation period. The addition of bovine serum albumin (10mg BSA/ml) to the dilution buffer resulted in no activity loss after a 3h incubation period (Figure 5.3). Additional experiments revealed 5mg BSA/ml was sufficient to maintain enzymic activity 3h after dilution.

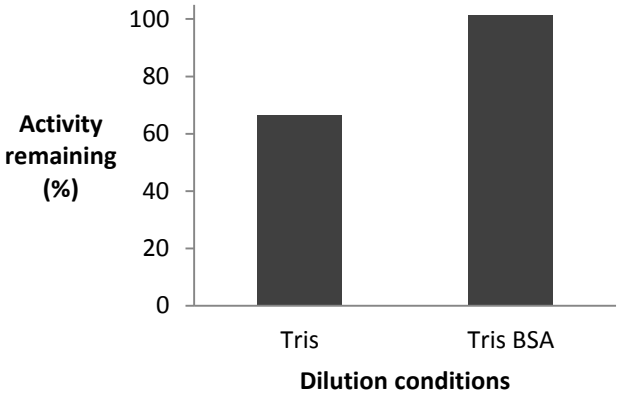


Figure 5.3 Relative acetate kinase activity remaining 3h post sample dilution. Purified recombinant AK activity was measured immediately after dilution (defined as 100%), and the remaining activity is shown after a 3h incubation period. Samples were kept on ice in 50mM Tris-HCl, pH 8.0, with or without 10mg BSA/ml of buffer.

5.3.3.2 Kinetic properties

Kinetic characterisation of recombinant AK was determined using the method described in Section 2.7.7.4. The purified protein was diluted in a 50mM Tris buffer, pH 8.0, containing 5mg BSA/ml, and the ionic strength of the assay was maintained at 0.1M with the addition of NaCl. Apparent substrate inhibition was observed with ADP; data were fitted to the substrate inhibition rate equation using SigmaPlot (Figure 5.4 A and B) and yielded a $V_{\max}$ of 22,300 (± 800) U/mg, where U is µmol of ATP produced per minute, a $K_{\mathrm{m(ADP)}}$ of 0.72 (± 0.04) mM, and a K_{i} of 6.4 (± 0.8) mM. The rate of AK displayed typical Michaelis-Menten kinetics towards acetyl-phosphate concentration (Figure 5.4 C and D); data were analysed by SigmaPlot and yielded a $V_{\max}$ of 14,200 (± 500) U/mg and a $K_{\mathrm{m(acetyl\text{-}phosphate)}}$ of 0.34 (± 0.04) mM. A slight curve was evident in Figure 5.4 D suggesting substrate inhibition with acetyl-phosphate, therefore the data were also fitted to the substrate inhibition rate equation and yielded a $V_{\max}$ of 16,000 (± 1,300) U/mg, a $K_{\mathrm{m(acetyl\text{-}phosphate)}}$ of 0.43 (± 0.08) mM, and a K_{i} of 26 (± 16) mM.

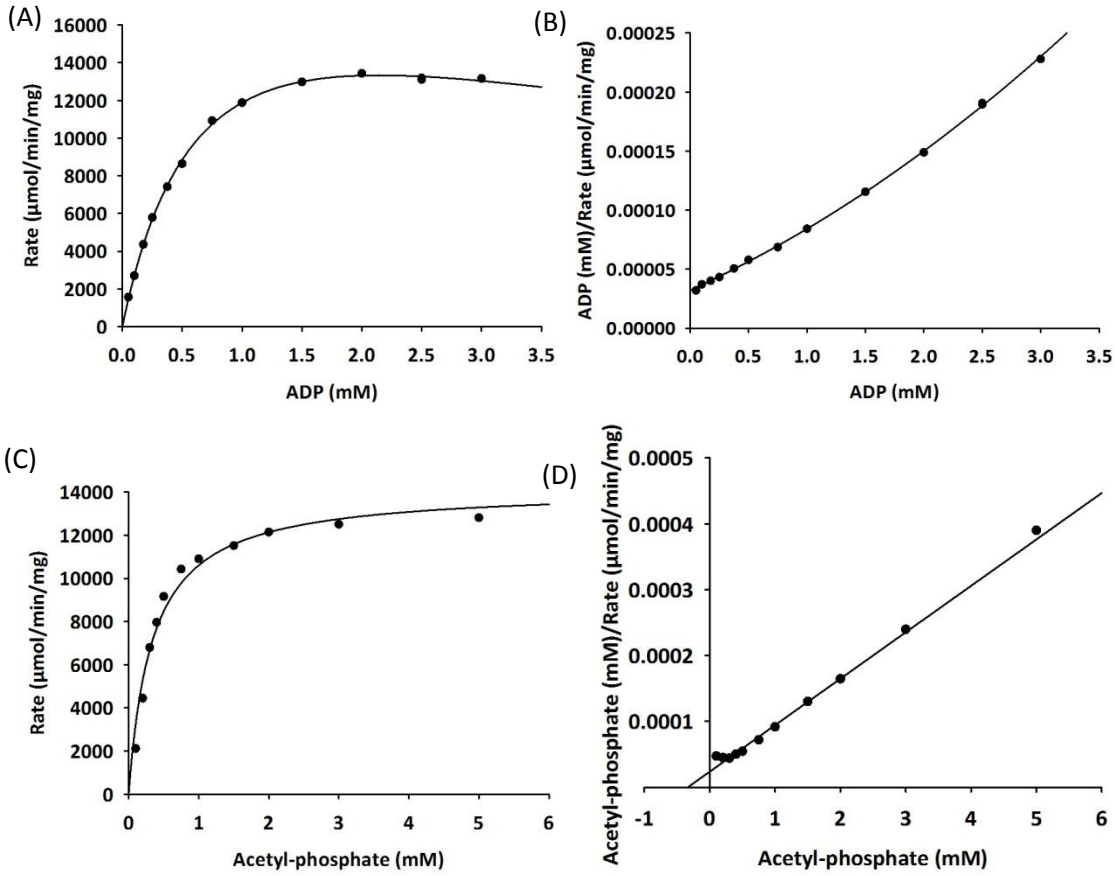


Figure 5.4 Acetate kinase activity. The relationship between specific activity (U/mg) and ADP concentration at a fixed acetyl-phosphate concentration of 5mM is displayed as (A) a Michaelis-Menten plot; and (B) a Hanes-Woolf plot. The relationship between specific activity (U/mg) and acetyl-phosphate concentration at a fixed ADP concentration of 3mM is displayed as (A) a Michaelis-Menten plot; and (B) a Hanes-Woolf plot.

5.4 Discussion

To date, neither phosphotransacetylase (PTA) nor acetate kinase (AK) have been characterised from *G. thermoglucosidasius*. Cloning of the *ak* gene, and expression of the recombinant enzyme in *E. coli*, provided a relatively easy system for the purification of AK for subsequent characterisation. The K_{m} values for the recombinant AK were similar to those determined in Chapter 3 for the native enzyme from *G. thermoglucosidasius* TM242 cell lysate. This suggests that the recombinant expression system produced an enzyme comparable to the natively expressed AK, and was a viable system for characterising the purified AK from *G. thermoglucosidasius* TM242.

The K_{m} values determined for both acetyl-phosphate and ADP were similar to those reported for *Methanosarcina thermophila* and *Clostridium acetobutylicum* (Ingram-Smith *et al.* 2005; Winzer *et al.* 1997). As observed in other organisms, substrate inhibition with ADP was evident (Kahane and Muhlrad 1979; Yoshimura 1978). However, it has not been

investigated whether this is an artefact of the enzyme assay. AK efficiently catalyses the reaction only if substrates ADP and ATP are bound to a divalent cation, in this case Mg^{2+}. The inhibition observed may have been true substrate inhibition, or due to a lack of Mg^{2+} ions present in the assay limiting the concentration of ADP bound to Mg^{2+} (Knorr *et al.* 2001).

The specific activity of the recombinant AK enzyme from *G. thermoglucosidasius* TM242 was significantly greater than that for other known AKs (BRENDA database). The specific activity was between 10- and 15-fold higher than the highest reported for AK on BRENDA, from *Geobacillus stearothermophilus* (formally *Bacillus stearothermophilus*) (Nakajima *et al.* 1978). The high activity and high expression levels (Chapter 3), in combination with the relatively low K_{m} values, suggest that *in vivo* acetyl-phosphate is rapidly converted to acetate by AK. This will be the case under the assumption that during fermentative growth on a sugar source ADP and acetyl-phosphate concentrations are high. Measuring the intracellular concentrations of these metabolites would improve the modelling of flux through this pathway.

To elucidate the physiological role of the PTA-AK pathway it would be advantageous to determine the kinetic parameters in the reverse direction, and to investigate potential inhibitors of these enzymes. However, due to the difficulties in using commercially sourced mesophilic enzymes in a coupled assay system to study the thermophilic AK this work was not achieved. Despite this, the data collected in this Chapter may suggest that *G. thermoglucosidasius* TM242 has evolved to catalyse the rapid conversion of acetyl-phosphate to acetate in order to store energy in the form of ATP. This would prevent the loss of energy that would result from the spontaneous hydrolysis of acetyl-phosphate, which would be enhanced at the elevated growing temperatures of *G. thermoglucosidasius*. Alternatively, AK could be used to control the levels of acetyl-phosphate, and hence control global regulation of the cell in response to environmental cues.

Chapter 6

Adaptation of the pTMO31 integration vector and deletion of *pta*

6.1 Introduction

The Gram-positive thermophile, *Geobacillus thermoglucosidasius*, utilises a wide range of substrates, including hexose and pentose sugars, as carbon and energy sources. The wild-type strain *G. thermoglucosidasius* NCIMB 11955 can ferment these substrates producing some ethanol but predominantly lactate. Cripps *et al.* (2009) demonstrated that this organism could be engineered for enhanced bioethanol production by disrupting the genes encoding lactate dehydrogenase and pyruvate formate lyase, and increasing the expression of the pyruvate dehydrogenase complex. The resulting strain, *G. thermoglucosidasius* TM242, can ferment the sugar substrates and produce ethanol at yields in excess of 90% of theoretical maximum. Acetate is the only other metabolite that is produced by *G. thermoglucosidasius* TM242 in significant quantities.

Acetate is an undesirable by-product because it can retard cellular growth and inhibit protein formation. Moreover, acetate production represents a diversion of carbon that might otherwise have generated biomass or ethanol. Acetate is generated by the activities of phosphotransacetylase (PTA) and acetate kinase (AK), which catalyse the conversion of acetyl-CoA and orthophosphate to acetate and ATP. The aim of this study was to disrupt the PTA – AK pathway by deleting the *pta* gene in the *G. thermoglucosidasius* TM242 strain.

Thermophilic organisms are, however, notoriously difficult to engineer at the genetic level for several reasons. Firstly, the transformation of thermophilic bacteria may be inhibited due to endospore formation, the cell envelope and the low permeability of the plasma membrane (Averhoff 2004; Culha *et al.* 2008; Silhavy *et al.* 2010). Secondly, the thermostability and efficiency of suitable marker genes and the replication origin of the plasmid DNA are essential for genetic manipulation of thermophiles (Lin and Xu 2013). Finally, the reported protocols for thermophilic transformations are limited (Taylor *et al.* 2009).

Cripps *et al*. (2009) developed a genetic toolkit and a transformation protocol to address some of these issues in order to create the TM242 strain from *G. thermoglucosidasius* 11955. It was reported that the integration vector created, pTM031, was capable of

autonomous replication in both *E. coli* at 37°C using selection with ampicillin (100µg/ml) and *G. thermoglucosidasius* at 55°C with selection on kanamycin (12.5µg/ml). The vector contained a temperature-sensitive replication origin that did not allow plasmid replication in *G. thermoglucosidasius* above 65°C (Figure 6.1). Any colonies growing above this temperature would therefore have integrated the complete vector into the genome. This system was capable of generating stable gene knockouts and gene replacements that were markerless, and hence did not require the continuous presence of antibiotics to maintain the phenotype.

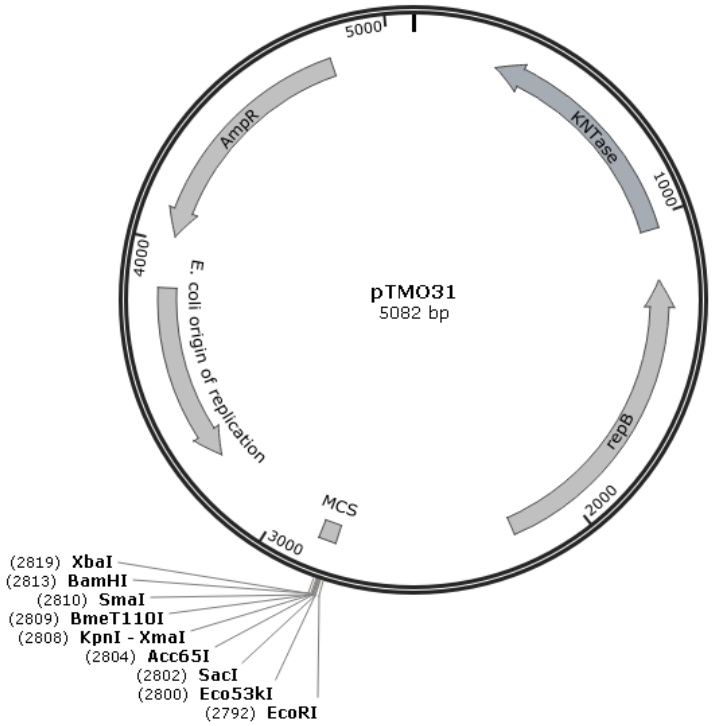


Figure 6.1 The backbone sequence of the *Geobacillus* integration vector, pTMO31. The construct contains an *E. coli* origin of replication, an MCS and the ampicillin resistance marker AmpR from pUC19. The repB is the thermolabile origin of replication and is responsible for plasmid DNA replication in *Geobacillus* species below 65°C. The repB and the kanamycin resistance marker, KNTase, are from pUB110. Chromosomal integration is achieved by insertion of sequences homologous to targeted regions on the chromosome.

However, the transformation efficiency of *G. thermoglucosidasius* with pTMO31 was between 5-20 transformants per µg of DNA. This is significantly lower than the efficiency with *Geobacillus* shuttle vectors pUCG18 (Taylor *et al.* 2008) and pUCG3.8 (Bartosiak-Jentys *et al.* 2013) of 1×10^4 and 2.8×10^5 tranformants per µg of DNA, respectively. The integration vector pTMO31 is a chimera of a general *E. coli* cloning vector pUC19 and the widely-used Gram-positive vector pUB110 from *Staphylococcus aureus*. Kanamycin nucleotidyltransferase (KNTase) activity is responsible for the kanamycin resistance trait of

pUB110. It has been demonstrated that pUB110 can confer kanamycin resistance to *Geobacillus stearothermophilus* (formally *Bacillus*) at temperatures below 55°C, but crucially thermostable mutants to the KNTase could increase transformation efficiency (Liao and Kanikula 1990). The most thermostable KNTase mutant isolated had two amino acid changes: Asp80-Tyr and Thr130-Lys.

This Chapter describes the adaptation of the integration vector to improve transformation efficiency, and the deletion of *pta*. The Δ*pta* and the parent TM242 strains are characterised both in terms of enzyme activity and their ability to grow fermentatively on hexose and pentose sugars. The pTMO31 and pUCG3.8 vectors were provided by TMO Renewables Ltd and Jeremy Bartosiak-Jentys, respectively. The bench-top fermentation and analysis were assisted by Joanne Neary and Matthew Lord at TMO Renewables Ltd.

6.2 Methods

6.2.1 Effects of kanamycin concentration on pTMO31 transformation

To compare the transformation efficiency of pTMO31 at differing concentrations of kanamycin, electrocompetent *G. thermoglucosidasius* TM242 cells were prepared as described in Section 2.6.9. Aliquots of cells (60µl) were transformed with 2µl of pTMO31 using the method described in Section 2.6.10. Transformants were recovered in 2SPY at 52°C for 4h in a shaking incubator set to 250rpm. Recovered cells were harvested and spread on TS agar plates containing kanamycin at concentrations between 1.2µg/ml and 12.5µg/ml and incubated at 52°C overnight. Colony forming units on each plate were observed. As a control, untransformed electrocompetent *G. thermoglucosidasius* TM242 aliquots were spread on TS agar plates containing the same concentrations of kanamycin.

6.2.2 Creation of pUB31

The gene in pTMO31 encoding the KNTase was replaced with the gene from pUCG3.8 that encodes a thermostable KNTase. This was achieved by digesting both plasmids with *Bsr*GI and *Bbs*I. The 789bp fragment excised from pUCG3.8 was ligated into the 4,293bp fragment of pTMO31, and the plasmid was transformed into *E. coli* JM109. The resulting plasmid, pUB31, was fully sequenced to confirm the presence of the new KNTase encoding gene.

6.2.3 Transformation efficiencies

To compare the transformation efficiency of pUB31 to that of pTMO31, fresh electrocompetent *G. thermoglucosidasius* TM242 cells were prepared as described in Section 2.6.9. Aliquots of cells (60µl) were transformed with 1µg of each vector using the method described in Section 2.6.10. Transformants were recovered in 2SPY at 52°C for either 2h or 4h in a shaking incubator set to 250rpm, prior to being spread on TS agar plates containing 12.5µg/ml kanamycin. Plates were incubated at the desired temperature for 16h. Colony forming units per µg of DNA were calculated.

6.2.4 Creation of a *pta* knockout construct

The *pta* gene and approximately 800bp sequences flanking either side of the gene were PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA, using primers ptaKOF and ptaKOR (Table 6.1). The PCR product was purified and digested with *Kpn*I and *Xba*I, prior to ligation into pUC19 which had been linearized with *Kpn*I and *Xba*I. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. Divergent primers, ptaKOdivF and ptaKOdivR (Table 6.1), were used to remove the *pta* gene while amplifying the pUC19 backbone and 800bp flanking regions. Purified PCR product was digested with *Not*I, and circularised by self-ligation. Ligated plasmids were replicated by transforming *E. coli* JM109 cells.

The pUC19 backbone now contained a 1,597bp insert that consisted of the DNA sequence that flanks the *pta* gene in the *G. thermoglucosidasius* TM242 genome. This was excised from pUC19 by digestion with *Kpn*I and *Xba*I, and ligated into pUB31 that had been linearized with the same restriction enzymes. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies that had grown on LB plates with carbenicillin.

Table 6.1 PCR primers used for creation of *pta* knockout construct. Restriction sites are underlined.

Name	Sequence (5'-3')
ptaKOF	GGGGTACCCGGGCGCTCACTTCGTCAAAACGCAT
ptaKOR	GCTCTAGAGTGCGAAATAATAAAAGTAAGACGGA
ptaKOdivF	AAGGAAAAAAGCGGCCGCTAATTCAGAAAAAGGTGACGAAACGT
ptaKOdivR	TTTTCCTTTTGCGGCCGCAAAAAATCGATTTATTAAGACAATTC

6.2.5 Isolation of mutant strains with *pta* deletion

Mutant strains with a deletion to the *pta* gene were isolated following the protocol described in Section 2.6.16.

6.2.6 Enzyme assays

Geobacillus cells were grown aerobically in 2SPYNG in 250ml baffled flasks. A cell lysate was prepared as described in Section 2.7.3 and enzyme assays were performed as described in Section 2.7.7.3.

6.2.7 Tube fermentations

Micro-aerobic tube fermentation cultures were carried out as described in Section 2.5.1. *Geobacillus* strains were revived from glycerol stocks and grown in 2SPYNG to use as a seed culture. USM medium containing 2% (w/v) yeast extract and either 2% (w/v) glucose or 2% (w/v) xylose was inoculated with the corresponding strain once an OD_{600} of 3 had been reached in the seed culture. The cell pellet was harvested and the supernatant analysed by HPLC after the culture had been incubated for 48h at 60°C.

6.2.8 Bench-top fermentations

Batch fermentations of *Geobacillus* strains were carried out in USM medium containing 2% (w/v) yeast extract and either 4% (w/v) glucose, 3% (w/v) xylose or 4% (w/v) mixed sugar [2.4% (w/v) glucose and 1.6% (w/v) xylose] as described in Section 2.5.2. Fermentations and subsequent HPLC analysis were performed at TMO Renewables Ltd in at least duplicate runs for each strain.

6.3 Results

6.3.1 Adaptation of the integration vector

6.3.1.1 Effect of kanamycin concentration on transformation efficiency with pTMO31

The effect of kanamycin concentration in TS agar plates on transformation efficiency with the pTMO31 integration plasmid was studied. Aliquots of *G. thermoglucosidasius* TM242 were either transformed with pTMO31 or left untransformed as the control. These cells were cultured on TS agar plates with kanamycin concentrations ranging from 1.2µg/ml to 12.5µg/ml. Despite Cripps *et al.* (2009) reporting that the pTMO31 transformants can be selected on kanamycin at 12.5µg/ml, no colonies were observed on plates with kanamycin concentrations of 3.2µg/ml and above. TS agar plates with 1.2µg/ml had 137 colonies where cells had been transformed, but there were 49 colonies on the control plate (Table 6.2). This suggests that this concentration is not sufficient to prevent the growth of *G. thermoglucosidasius*.

Table 6.2 The effect of kanamycin concentration on transformation efficiency. *G. thermoglucosidasius* TM242 cells, either untransformed or transformed with pTMO31, were cultured on TS agar plates containing kanamycin, and the number of colony forming units were observed.

	Colony forming units (colonies per plate)					
Kanamycin (µg/ml)	1.2	3.2	5.2	7.2	9.2	12.5
pTMO31	137	0	0	0	0	0
Untransformed	49	0	0	0	0	0

6.3.1.2 Creation of pUB31

The results from Section 6.3.1.1 suggested that the lack of kanamycin resistance could be limiting transformation efficiencies with the pTMO31 vector. Kanamycin resistance from pTMO31 is conferred by the kanamycin nucleotidyltransferase (KNTase), the gene for which originally came from a *Staphylococcus aureus* plasmid, pUB110 (Cripps *et al.* 2009; McKenzie *et al.* 1986). The work by Liao and Kanikula (1990) revealed that kanamycin resistance in *Geobacillus stearothermophilus* could be enhanced by using mutant KNTases that were more thermostable than the original expressed from pUB110. Thermostability was increased when mutants were isolated containing two amino acid changes; Asp80-Tyr and Thr130-Lys.

The thermostable KNTase encoding gene described by Liao and Kanikula (1990) was incorporated into a *Geobacillus* spp.-*E. coli* shuttle vector (pUCG18) by Taylor *et al.* (2008). The gene encoding the KNTase in pTMO31 was replaced by digesting pTMO31 and pUCG3.8 (a variant of pUCG18) with *Bsr*GI and *Bbs*I. This resulted in the integration vector now containing a gene that encodes a thermostable KNTase. This plasmid was named pUB31.

6.3.1.3 Comparison of transformation efficiencies between pTMO31 and pUB31

The transformation efficiency of pUB31 and pTMO31 were compared by transforming electrocompetent *G. thermoglucosidasius* TM242 cells with 1µg of each vector. Cells were recovered prior to spreading on TS agar plates containing 12.5µg/ml kanamycin. After a 16h incubation period the colony forming units were calculated. No colonies grew of cells transformed with pTMO31; however, cells transformed with pUB31 were able to grow at temperatures below 64°C (Figure 6.2).

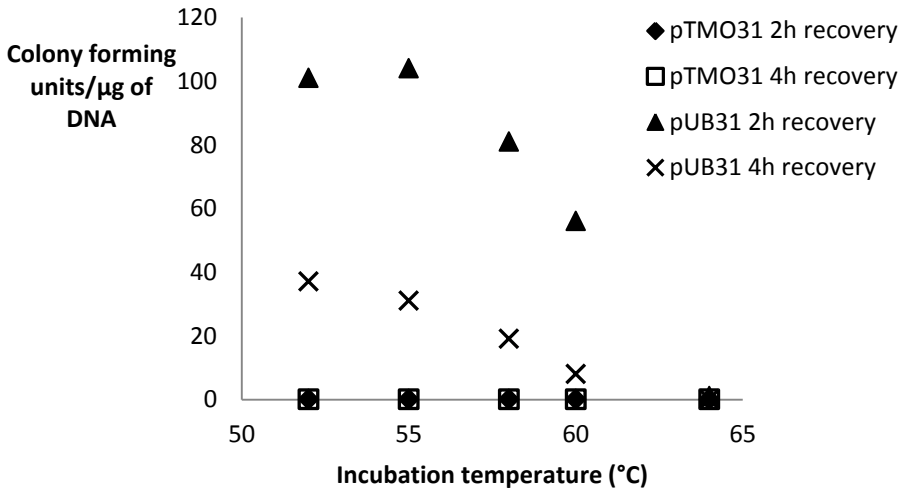


Figure 6.2 Comparing transformation efficiencies of *G. thermoglucosidasius* TM242 with either pTMO31 or pUB31. Transformed cells were allowed to recover for either 2h or 4h at 55°C in a shaking incubator prior to being spread onto TS agar plates containing 12.5µg/ml kanamycin. Plates were incubated at the desired temperature for 16h before colony forming units were counted.

6.3.2 Deletion of *pta* and strain characterisation

6.3.2.1 Creation of the *pta* knockout construct

The construct designed to delete the *pta* gene from *G. thermoglucosidasius* TM242 was created in three stages. Initially, the *pta* gene and DNA sequences flanking either side of the gene were PCR-amplified from genomic DNA and ligated into the *E. coli* cloning vector,

pUC19. The *pta* gene was deleted from this construct by divergent PCR, and subsequent self-ligation left an insert sequence that was homologous to the regions flanking the *pta* gene. This insert was then ligated into the pUB31 multiple cloning site, and the construct was replicated in *E. coli* JM109 cells.

The *pta*KO-pUB31 construct contains sequences that are homologous to the DNA regions that flank the *pta* gene in *G. thermoglucosidasius* TM242, which provides the basis for gene deletion (Figure 6.3). The construct also contains the thermolabile origin of replication, and genes responsible for conveying ampicillin resistance in *E. coli* and kanamycin resistance in *Geobacillus* species.

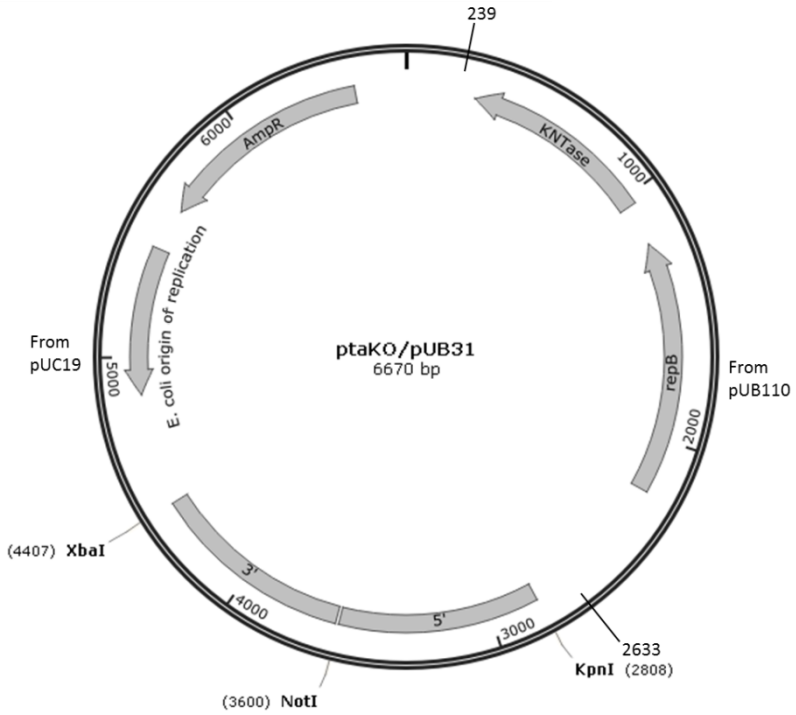


Figure 6.3 The construct designed to delete *pta* from *G. thermoglucosidasius*. The construct contains an *E. coli* origin of replication and the ampicillin resistance marker AmpR from pUC19. The repB is the thermolabile origin of replication and is responsible for plasmid DNA replication in *Geobacillus* species below 65°C. The repB and the kanamycin resistance marker, KNTase, are from pUB110. The region highlighted as 5' and 3' are homologous to sequences that are 5' and 3' to *pta* in *G. thermoglucosidasius*, respectively.

6.3.2.2 Creation of the *pta* knockout strain

The *pta* in *G. thermoglucosidasius* TM242 strain was targeted for gene knockout, by transforming cells with the *pta*KO-pUB31 construct. Strains in which the *pta* had been deleted were isolated following the procedure explained in Figure 6.4. Kanamycin-sensitive, double-crossover isolates were identified by colony PCR as either successful gene knockouts

or having reverted back to the parent TM242 strain (Figure 6.5). Of the kanamycin sensitive colonies screened, only 6% were identified as successful gene knockouts, suggesting that the deletion of *pta* may be somewhat detrimental to the organism's survival.

Enzymic assays with cell lysates from aerobically grown *G. thermoglucosidasius* confirmed the deletion of *pta*. The Δ*pta* cell lysate had 8% of the activity with acetyl-CoA and orthophosphate compared to the TM242 strain (Figure 6.6). The activity with propionyl-CoA was also reduced in the Δ*pta* strain, while the activity with butyryl-CoA remained relatively consistent. This could be attributed to effects on altering expression of alternative phosphotransacylase enzymes, or due to the direct loss of PTA which has promiscuous activity with propionyl-CoA (see Section 4.3.7 for details).

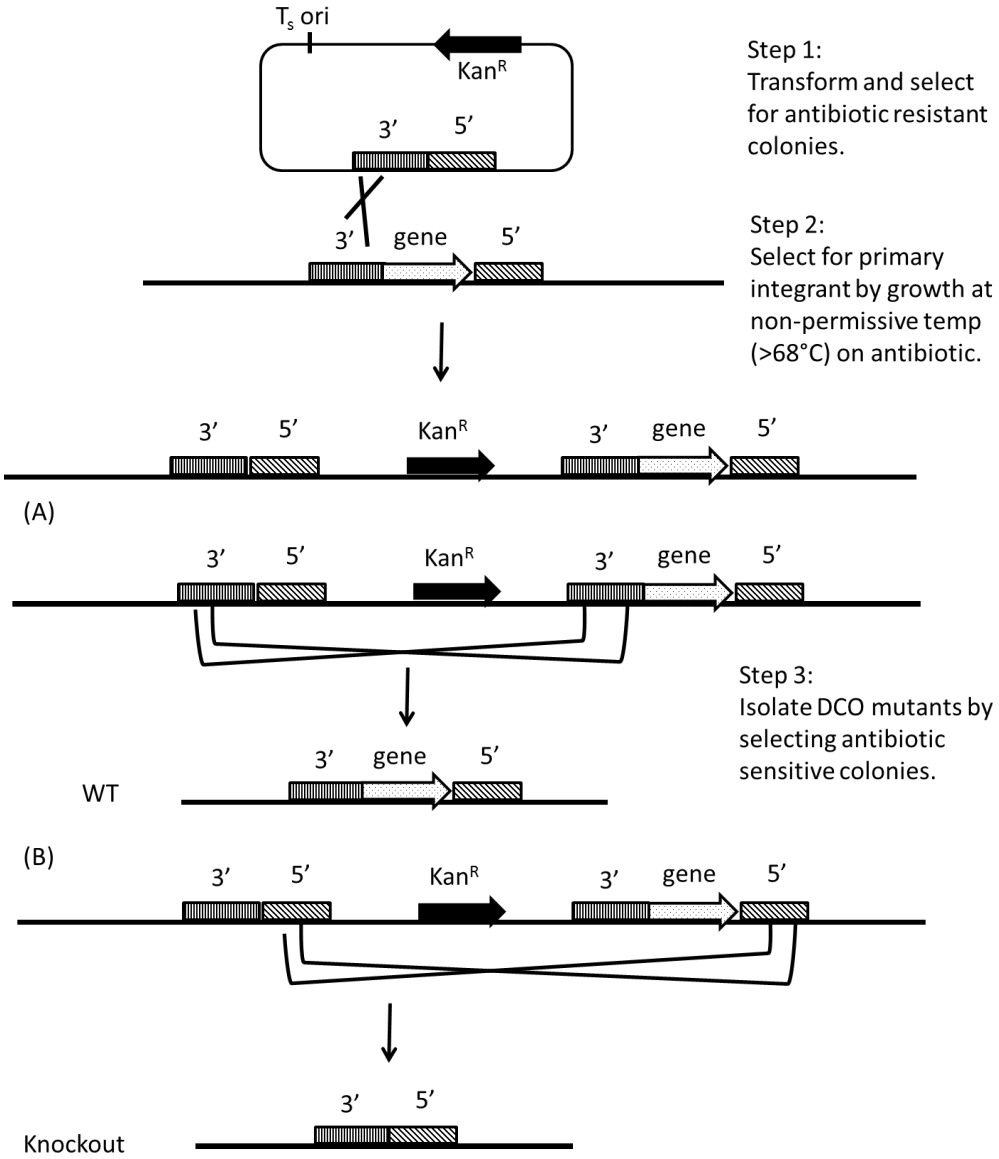


Figure 6.4 Gene knockout strategy using pUB31-based constructs. The knockout construct contains an antibiotic resistance marker, a temperature sensitive replication of origin and two sequences that are homologous to regions of DNA 5′- and 3′- of the gene targeted for deletion. After kanamycin resistant colonies are isolated (Step 1), primary integrants, which have incorporated the construct into the chromosomal DNA, are isolated by selecting colonies that have grown at 68°C on kanamycin (Step 2). Double crossover (DCO) mutants are selected for kanamycin sensitivity after a period of subculturing; this second recombination event will either generate the parent/wild-type strain (A) or the knockout mutant (B).

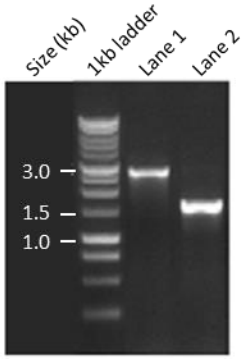


Figure 6.5 PCR colony screen of kanamycin sensitive double crossover mutants. Kanamycin sensitive colonies were screened by PCR with primers that anneal to the regions of DNA flanking the *pta*. Lane 1 shows an amplicon of approximately 2.7kb, which is the expected size of the amplicon from the *G. thermoglucosidasius* TM242 strain with an intact *pta*. Lane 2 shows a band of approximately 1.6kb, which is the expected size of the amplicon for a *pta* deleted strain.

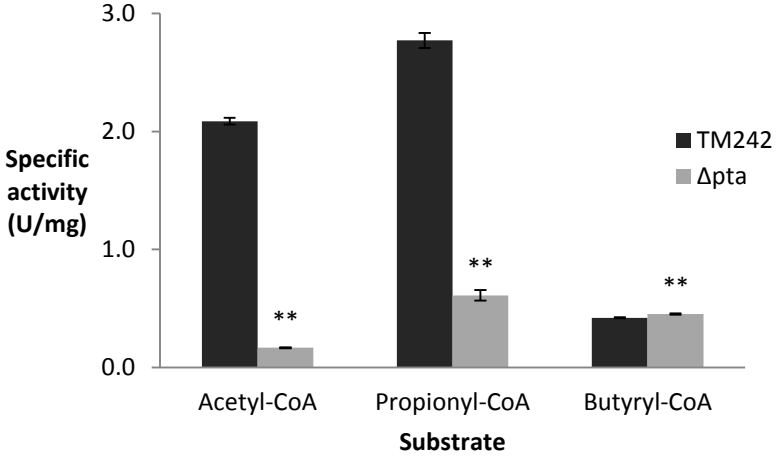


Figure 6.6 Comparing phosphotransacylase activity in cell lysate of *G thermoglucosidasius* TM242 and Δ*pta*. Cell lysates from aerobically grown cultures were assayed with 0.1mM acyl-CoA and 50mM phosphate. Each assay was performed in triplicate and the relevant background rate was subtracted. The student's t-test was used to determine the significance of any differences between the rate of activity from TM242 cell lysate compared to the activity from Δ*pta* cell lysate (one star indicates a p-value <0.05, two stars indicate a p-value <0.01 at 95% confidence limits).

6.3.2.3 Strain characterisation

To analyse the effects of the *pta* deletion from the TM242 strain, the fermentation products were analysed from both micro-aerobic tube cultures (Figure 6.7), and from batch fermentations in a bench-top fermenter (Figure 6.8, Figure 6.9 and Figure 6.10). The tube cultures indicate that the Δ*pta* strain has a similar phenotype to TM242 when these strains are grown on a USM medium containing glucose. A 10% reduction in ethanol yield was observed with the Δ*pta* strain grown on xylose, compared to TM242. Succinate, pyruvate and lactate yields were also higher under growth on xylose. The Δ*pta* strain still produced

acetate to similar levels as TM242 under xylose growth; however, a decrease of 17% was observed when the medium contained glucose.

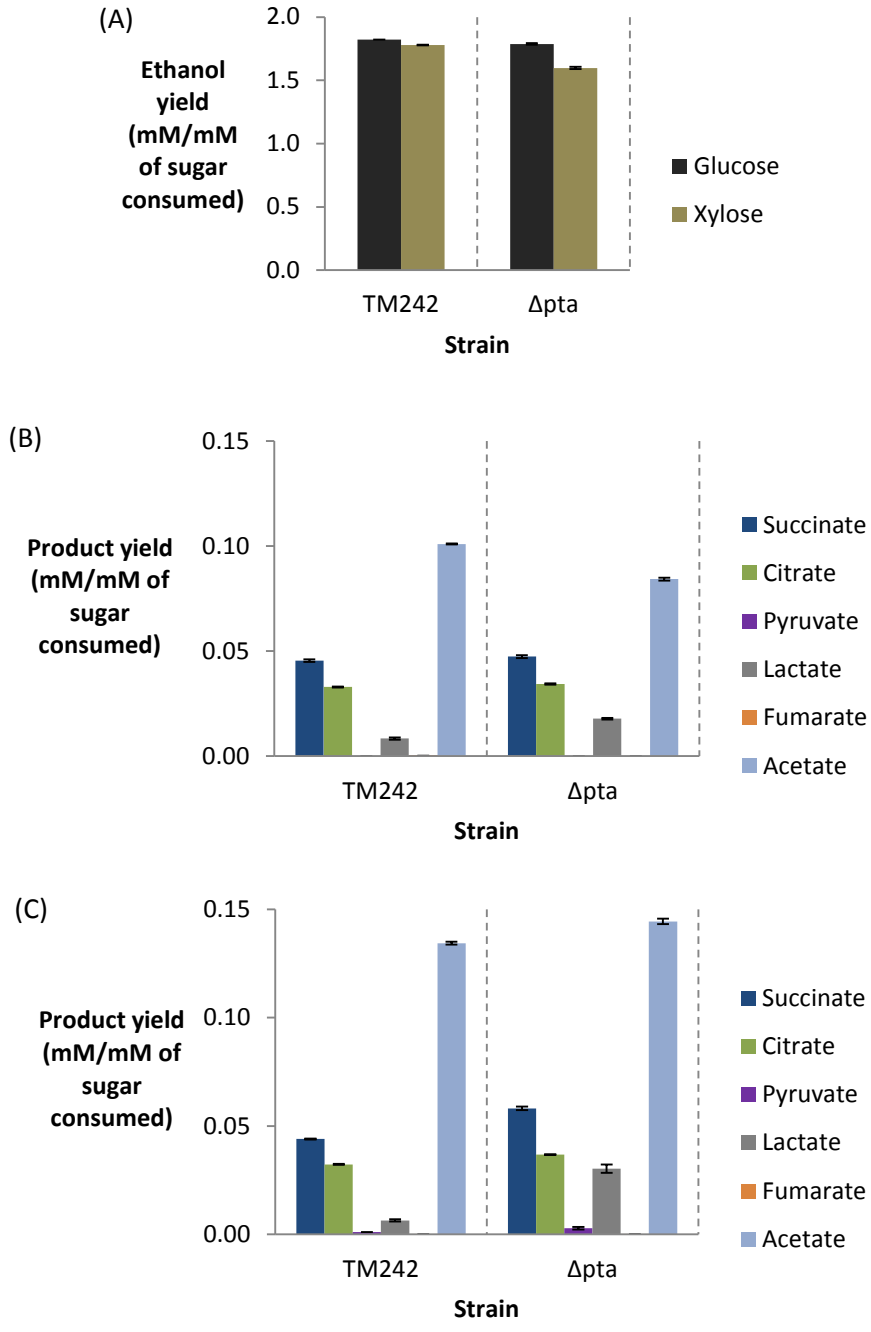


Figure 6.7 End-point products of fermentation by *G. thermoglucosidasius*. Metabolites were determined from tube fermentation cultures of TM242 and Δ*pta* grown on USM containing 2% (w/v) of either glucose or xylose. Cultures were grown at 60°C, shaking for 48h in sealed falcon tubes to induce fermentative metabolism. Products are given as a yield in terms of mM of product per mM of sugar consumed. Bar charts represent the ethanol yields on both sugars (A), and the organic acid profiles for glucose grown cells (B) and xylose grown cells (C).

Tube cultures are a useful method of rapidly analysing fermentation products, although the culture is sampled at only one time-point. Therefore, further characterisation was conducted in a bench-top fermenter where cells were grown on a USM medium containing 4% (w/v) glucose (Figure 6.8). The results indicate that the Δ*pta* strain consumed glucose more rapidly, and grew at a faster rate, than TM242. The ethanol concentration reached its peak 4.6h sooner for Δ*pta* culture, and was 10% higher than that for TM242. The acetate concentration was almost 50% lower at the time of peak ethanol production for the Δ*pta* strain compared to TM242.

The breakdown of lignocellulosic material results in the release of significant quantities of pentose sugars, notably xylose. *G. thermoglucosidasius* is capable of metabolising pentose sugars, and therefore the fermentation profile of Δ*pta* and TM242 were compared on a USM medium containing xylose (Figure 6.9). Growth of the TM242 strain ceased after the aeration and the agitation were decreased (t = 4h, Figure 6.9 (A)); as a result, the culture had not consumed all the available xylose by the 14h time-point. *G. thermoglucosidasius* TM242 is capable of the co-consumption of glucose and xylose (Cripps *et al.* 2009). To determine whether the Δ*pta* strain retains this capability, batch fermentations were run on USM medium containing a total of 4% mixed sugar (glucose + xylose); the results can be seen in Figure 6.10.

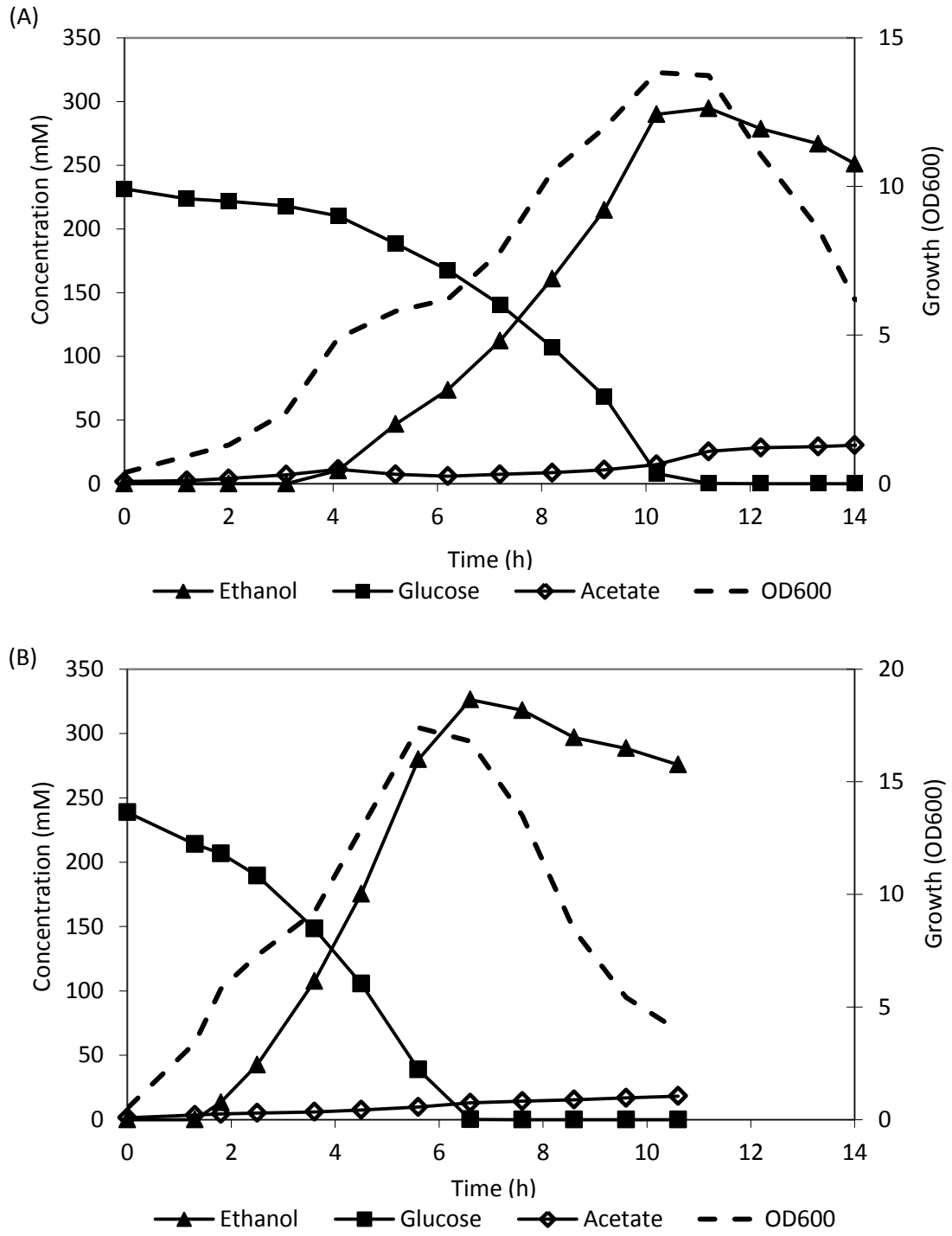


Figure 6.8 Products of fermentation by *G thermoglucosidasius*. Cells were cultured on 4% (w/v) glucose in a USM medium containing 2% (w/v) yeast extract as described in Section 2.5.2. The x-axis displays the time post inoculation, the primary y-axis displays the concentration of metabolites, and the secondary y-axis displays the growth as OD_{600}. The anaerobic switch was performed at 4h for TM242 (A) and 2h for Δ*pta* (B).

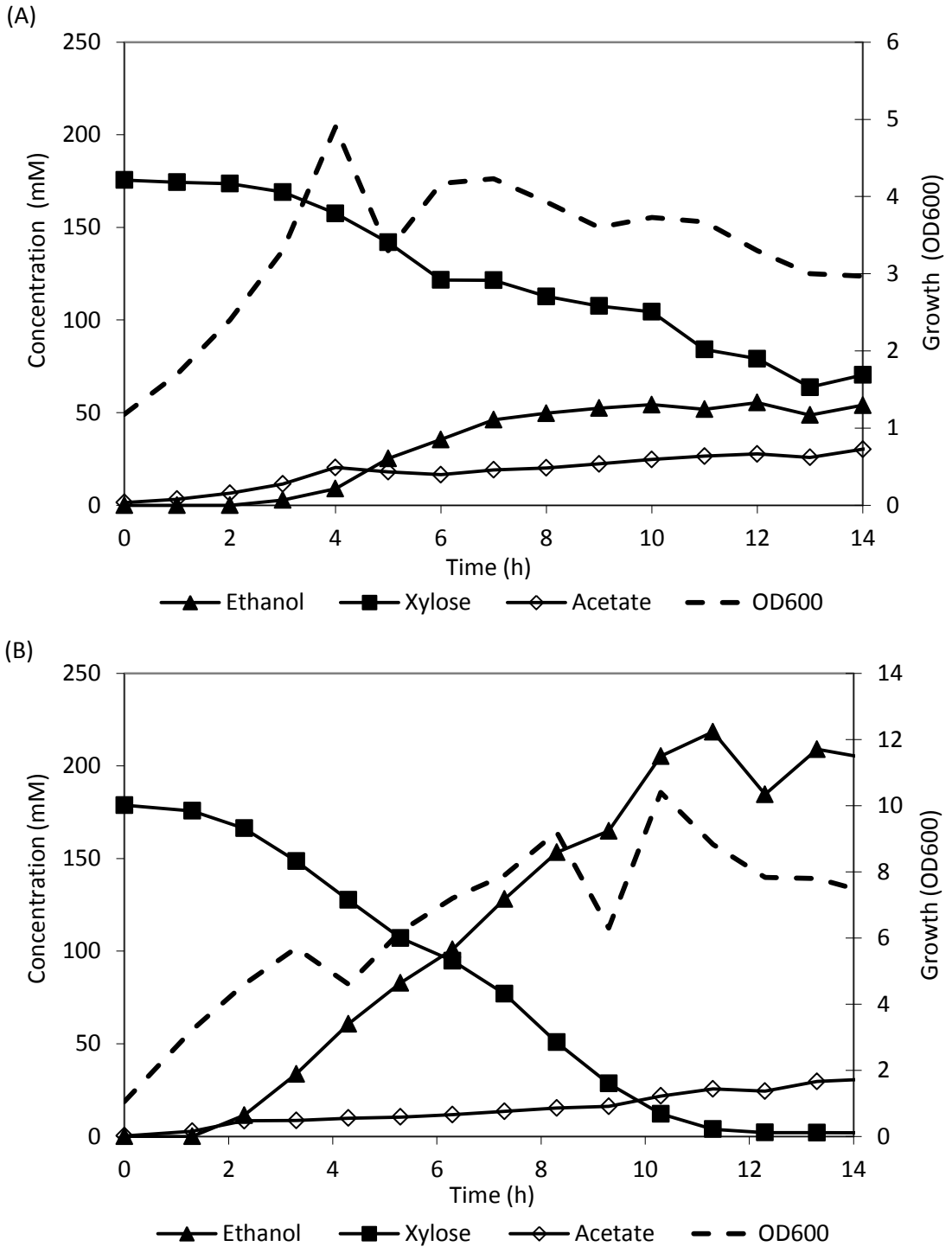


Figure 6.9 Products of fermentation by *G thermoglucosidasius*. Cells were cultured on 3% (w/v) xylose in a USM medium containing 2% (w/v) yeast extract as described in Section 2.5.2. The x-axis displays the time post inoculation, the primary y-axis displays the concentration of metabolites, and the secondary y-axis displays the growth as OD_{600}. The anaerobic switch was performed at 4h for TM242 (A) and 3.3h for Δ*pta* (B).

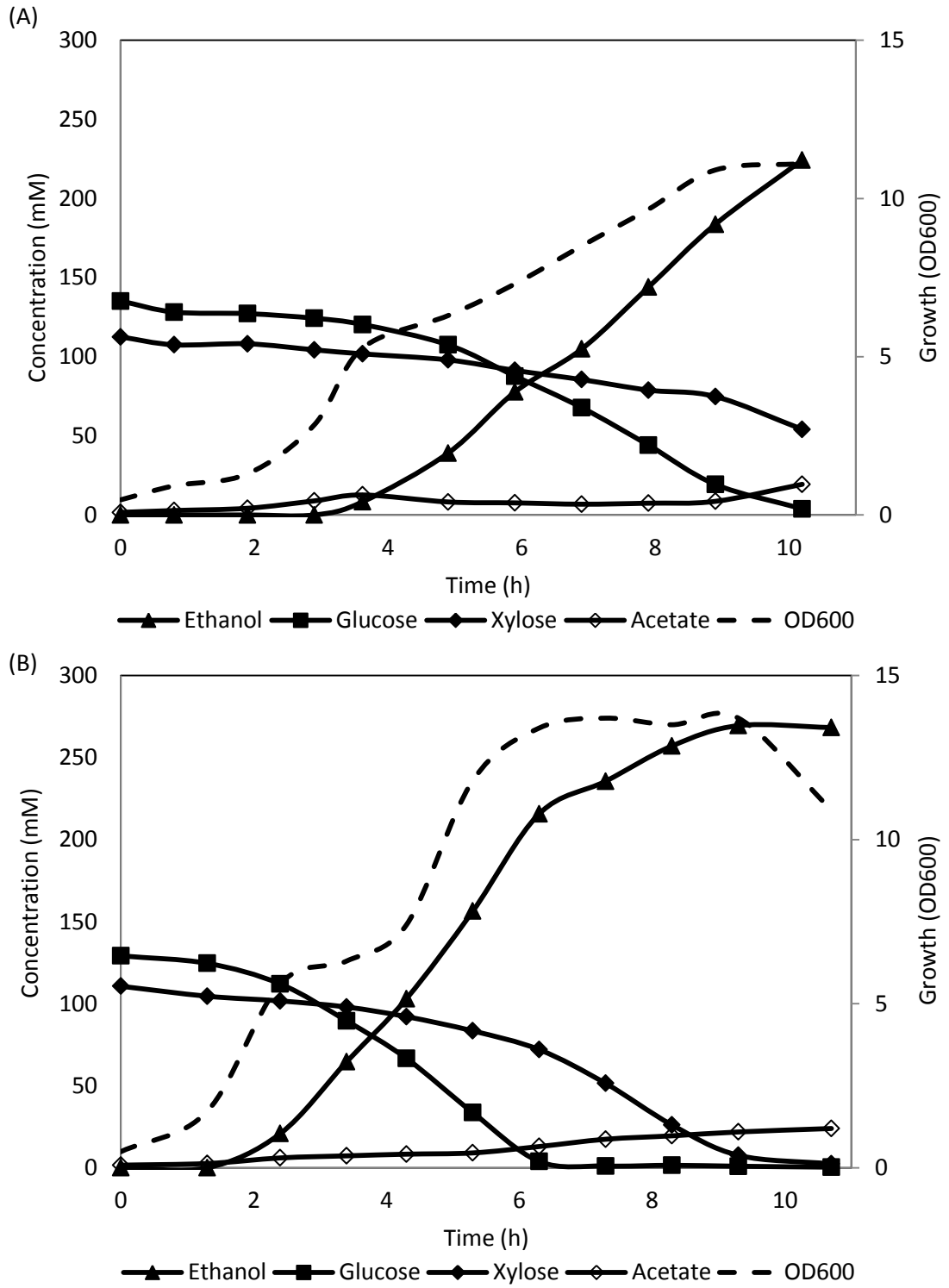


Figure 6.10 Products of fermentation by *G thermoglucosidasius*. Cells were cultured on 2.4% (w/v) glucose and 1.6% (w/v) xylose in a USM medium containing 2% (w/v) yeast extract as described in Section 2.5.2. The x-axis displays the time post inoculation, the primary y-axis displays the concentration of metabolites, and the secondary y-axis displays the growth as OD_{600nm}. The anaerobic switch was performed at 3.6h for TM242 (A) and 2.4h for Δ*pta* (B).

6.4 Discussion

The introduction of a gene encoding a thermostable kanamycin nucleotidyltransferase (KNTase) into the integration vector improved the transformation efficiency of *G. thermoglucosidasius* significantly. This vector improves the toolkit available for genetic manipulation of *Geobacillus* species and has been used to delete the *pta* gene from the *G. thermoglucosidasius* TM242 strain in this study.

Despite this improvement, the genetic toolkit available for *Geobacillus* species remains limited in comparison to mesophilic model organisms, and indeed to other thermophiles. To increase the speed and efficiency of isolating double-crossover mutants with the desired genotype, a positive-selection system would be required. Such systems have been developed for use in *Clostridium thermocellum* (Tripathi *et al.* 2010) and *Geobacillus kaustophilus* (Suzuki *et al.* 2012) based on strains auxotrophic for uracil. These approaches require the initial disruption of *pyrF*, which creates a strain that requires uracil supplements to survive. However, this may not be a viable approach to take when creating a strain for commercial use.

The deletion of *pta* from *G. thermoglucosidasius* TM242 resulted in a strain that still generated acetate, suggesting an alternative pathway must be responsible for the acetate produced in this strain. The PTA has been a target in other organisms, including *E. coli* and *Bacillus subtilis*, which similarly has resulted in strains that still produce acetate, but in lower amounts (Presecan-Siedel *et al.* 1999; Wang *et al.* 2012; 2013). To the author's knowledge, no-one has identified which pathway is responsible for the acetate produced in the PTA null strains. However, suggestions have included pyruvate oxidase, aldehyde dehydrogenase, alternative phosphotransacylases or an unidentified pathway.

The analysis of the fermentation profiles indicated that the Δ*pta* strain grew at a faster rate than TM242 and produced ethanol to a higher concentration. The peak ethanol for the Δ*pta* strain was at a yield of 0.42 and 0.38 g/g of sugar consumed for batch fermentations on glucose and xylose, respectively (Table 6.3). The ethanol yield was significantly greater than that observed for TM242 fermentation of xylose in particular. However, the reported ethanol yields for TM242 by Cripps *et al*. (2009) were 0.42 and 0.35 g/g of sugar consumed for glucose and xylose, respectively. The growth conditions used by this group were similar,

but not identical to the ones used here, and may account for why the Cripps group did not observe a reduction in growth post micro-aerobic switch for the TM242 strain on xylose.

At the time of peak ethanol concentration from glucose fermentations, the metabolite profile of the Δ*pta* strain is similar to TM242 (Table 6.3). The only significant difference is that the acetate concentration is almost 50% lower in the Δ*pta* fermentation, suggesting that the PTA is involved in producing a significant amount of acetate in the TM242 strain. Due to the discrepancies observed for TM242 xylose fermentations, the Δ*pta* profile was compared against the TM242 data from this study and that of Cripps *et al*. (2009). When compared to the Cripps *et al*. (2009) TM242 data the Δ*pta* strain has an increased ethanol yield, decreased acetate yield and an increased pyruvate yield at the time of peak ethanol production. This suggests that, without a functional PTA, the carbon flux can be diverted away from acetate towards ethanol production, but also results in the accumulation and secretion of pyruvate, possibly due to inhibition of the pyruvate dehydrogenase complex by acetyl-CoA. In addition, the Δ*pta* strain produced less lactate compared to the TM242 strain, particularly on xylose growth. Considering both strains have a disrupted *ldh* gene which encodes an L-LDH, the lactate may be produced by a D-LDH.

Table 6.3 Metabolite profiles of *G. thermoglucosidasius* batch fermentations. Fermentations of TM242 (Δ*ldh, Δpfl*, *pdh*^{up}) and Δ*pta* (Δ*ldh, Δpfl*, *pdh*^{up}, Δ*pta*) were carried out in USM medium containing 2% (w/v) yeast extract and indicated concentrations of substrate. The table details the concentration of metabolites produced and the residual substrate concentration at the time of peak ethanol levels. The maximum ethanol level gives the time post-inoculation that peak ethanol concentration was observed. It is worth noting that the TM242 fermentation of xylose did not agree with the observations of Cripps *et al*. (2009), and that the data set recorded for TM242 on mixed sugar did not reach the peak ethanol concentration.

Strain	Substrate	Max ethanol level (h)	Metabolite concentration (mM)							Ethanol yield (g/g)
			Initial substrate	Residual substrate	Ethanol	Acetate	Lactate	Succinate	Pyruvate	
TM242	Glucose	11	231	0	295	25	5	14	1.5	0.39
TM242	Xylose	12	176	79	55	28	20	6	15	0.18
TM242	Mixed sugar Glucose Mixed sugar Xylose	10	135 113	4 54	224	19	9	10	10	0.36
Δ*pta*	Glucose	6.6	239	0	326	13	3	12	1.5	0.42
Δ*pta*	Xylose	11	179	4	218	26	9	12	3.9	0.38
Δ*pta*	Mixed sugar Glucose Mixed sugar Xylose	9.3	129 111	1 8	269	22	7	15	2	0.36

In conclusion, this Chapter has detailed the adaptation of the integration vector to include a gene encoding a thermostable kanamycin nucleotidyltransferase, which has resulted in increased transformation efficiency of *Geobacillus* species. This vector has been used to

delete the gene encoding phosphotransacetylase from the TM242 (Δ*ldh, Δpfl*, *pdh*^{up}) strain. This has resulted in a strain that grows faster than the current TM242 strain, and produces ethanol at a faster rate, which may prove to be a potentially useful candidate in a commercial process. However, the Δ*pta* strain still produces acetate, despite a reduction of PTA-specific activity of greater than 92%. How this acetate is still produced remains a focus for this Thesis.

Chapter 7
Could a promiscuous phosphotransacylase compensate for the loss of phosphotransacetylase?

7.1 Introduction

The deletion of *pta* from the *G. thermoglucosidasius* TM242 strain did not abolish acetate production. Similar observations have been seen in PTA null mutants of other bacteria including *Bacillus subtilis* (Presecan-Siedel *et al.* 1999; Wang *et al.* 2012), *E. coli* (Chang *et al.* 1999; Wang *et al.* 2013), *Salmonella typhimurium* (Levine *et al.* 1980), and *Clostridium* species (Green *et al.* 1996; Zhu *et al.* 2005). The common theme in all these studies is that PTA-specific activity of the mutant decreased, although acetate production was not eliminated suggesting alternative pathway(s) must be present. Another major acetate-producing pathway known in *E. coli* is via pyruvate oxidase (PoxB), which catalyses the decarboxylation of pyruvate to form acetate and CO_2. Other suggestions of alternative acetate production pathways have included phosphoketolase, which catalyses the reaction between xylulose-5-phosphate and phosphate to form acetyl-phosphate and glyceraldehyde-3-phosphate. The acetyl-phosphate generated could subsequently be converted to acetate via acetate kinase (AK). However BLAST analysis of the *G. thermoglucosidasius* TM242 genome did not reveal any hits for either PoxB or phosphoketolase.

The presence of an additional phosphotransacylase (PTAC) in *G. thermoglucosidasius* with catalytic activity with acetyl-CoA could substitute for the loss of PTA in the Δ*pta* strain. The acetyl-phosphate generated by this PTAC could subsequently be converted to acetate by AK. The observation that 8% of PTA-specific activity (acetyl-CoA + phosphate → acetyl-phosphate + CoA) remained in the cell lysate of the *G. thermoglucosidasius* Δ*pta* strain, compared to TM242 (Figure 6.6) suggests that an additional PTAC may be present and active with acetyl-CoA.

Searches of the *G. thermoglucosidasius* TM242 genome revealed two genes that encode additional putative PTAC enzymes. One enzyme was annotated as a protein similar to PduL, a phosphotranspropionylase (PTP) that catalyses the reversible transfer of the propionyl-group between CoA and orthophosphate (Equation 7.1). The second was annotated as

phosphotransbutyrylase (PTB), which catalyses the reversible transfer of the butyryl-group between CoA and orthophosphate (Equation 7.2).

$$CH_3CH_2COCoA + HPO_4^{2-} \rightleftarrows CH_3CH_2COOPO_3^{2-} + CoA \quad (7.1)$$

$$CH_3CH_2CH_2COCoA + HPO_4^{2-} \rightleftarrows CH_3CH_2CH_2COOPO_3^{2-} + CoA \quad (7.2)$$

PTAC enzymes from other bacteria have been reported to have a degree of substrate promiscuity, including the ability to use acetyl-CoA as a substrate. A PTP characterised from *Salmonella enterica* was found to accept both propionyl-phosphate and acetyl-phosphate as substrates (Liu *et al.* 2007). Measured in the acyl-CoA forming direction, they determined this PTP was catalytically more active with propionyl-phosphate than acetyl-phosphate. This correlates with its proposed function in *S. enterica*, where the PTP is expressed from *pduL*, a gene located within a cluster of 23 genes (*pdu*) involved in 1,2-propanediol utilisation. In addition, Liu *et al*. (2007) demonstrated that ectopic expression of PTP (named PduL in their study) could alleviate the growth defect of a Δ*pta* mutant on acetate minimal medium, suggesting PTP can compensate for the loss of PTA in *S. enterica*.

Fermentation of rhamnose and fucose, common sugars found in plant cell walls, results in the production of 1,2-propanediol (1,2-PD). The degradation of 1,2-PD is a complex process that involves the compartmentalisation of catabolic enzymes within carboxysomes (polyhedral body-microcompartments). But the ability to degrade 1,2-PD provides a selection advantage to microorganisms in anaerobic environments, including those in the depths of soil environments, a common place to find *Geobacilli* which have also been observed to produce carboxysomes (Steve Bowden, University of Bath, personal communication).

Similarly, phosphotransbutyrylase (PTB) from certain *Clostridium* species has been demonstrated to catalyse the transfer of the acetyl-group between CoA and orthophosphate (Thompson and Chen 1990; Wiesenborn *et al.* 1989). The PTB enzymes characterised in these studies appear to be highly promiscuous with the acyl-CoA substrates, and have been shown to use acetyl-CoA, n-propionyl-CoA, n-valeryl-CoA, isobutyryl-CoA and isovaleryl-CoA in addition to n-butyryl-CoA. In certain Clostridia, PTB combines with butyrate kinase (BK) to convert butyryl-CoA to butyrate, a pathway that is linked with the acetone-butanol-ethanol (ABE) fermentation process. ABE fermentation is

the process by which solvetogenic Clostridia ferment sugars to a mixture of acetone, butanol and ethanol by switching their metabolism from acid production to solvent production upon receiving environmental pH cues.

The promiscuous nature of the PTP and PTB enzymes discussed in the studies above provided the basis of the current Chapter. Genes encoding PTP, PTB and BK were cloned from *G. thermoglucosidasius* TM242. Characterisation of the recombinant enzymes demonstrated that *G. thermoglucosidasius* does possess promiscuous PTAC enzymes. It is worth noting that a gene encoding a protein analogous to propionate kinase was not discovered in the *G. thermoglucosidasius* genome. A series of genetic manipulations were made that indicated the PTB rather than PTP could be involved in acetate production in the Δ*pta* strain. Further strain characterisation demonstrated that despite the abolition of PTA-specific activity, acetate was still generated by these strains, suggesting alternative acetate-producing pathways must exist in *G. thermoglucosidasius*. It was also discovered that strains carrying deletions to *pta* no longer re-utilise acetate during the shift to the stationary phase, indicating that PTA is important for the assimilation of acetate in this organism. In this Chapter kinetic characterisation of PTB and enzymatic assays of BK were carried out by supervised undergraduate project student Eilidh Livingstone.

7.2 Methods

7.2.1 Cloning

Cloning of the *ptp* and *ptb* constructs used general molecular biology techniques described in Section 2.6. Vector-specific primers were used for DNA sequencing to determine the correct sequence had been cloned. pET-vector clones were transformed into *E. coli* BL21(DE3) cells for protein expression.

7.2.1.1 *ptp*-pET-28a(+)

The *ptp* gene identified by ERGO™ Integrated Genomics was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA, using primers *ptp*F and *ptp*R (Table 7.1). The PCR product was purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *ptp* gene was excised from the pGEM®-T vector by restriction digest with *Nhe*I-HF and *Xho*I, agarose gel purified and ligated into pET-28a(+) that had been

linearized with *Nhe*I-HF and *Xho*I. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the construct.

7.2.1.2 *ptb*-pET-28a(+)

The *ptb* gene identified by ERGO™ Integrated Genomics was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA, using primers *ptb*F and *ptb*R (Table 7.1). The pET-28a(+) vector was digested with *Nde*I. The digested vector and PCR amplicon were purified by agarose gel electrophoresis, and the construct was assembled using the Gibson isothermal enzymic DNA assembly method (Gibson *et al.* 2009). Constructs were transformed into *E. coli* JM109 cells to amplify.

7.2.1.3 *bk*-pET vector constructs

The *bk* gene identified by ERGO™ Integrated Genomics was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA, using primers *bk*F and *bk*R (Table 7.1). The PCR product was purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *bk* gene was excised from the pGEM®-T vector by restriction digest with *Nhe*I and *Xho*I, agarose gel purified and ligated into pET-28a(+) that had been linearized with *Nhe*I-HF and *Xho*I. The *bk* clone was also inserted into pET-24a(+) by digesting the *bk*-pET-28a(+) construct and pET-24a(+) with *Nhe*I and *Xho*I. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the construct.

Table 7.1 PCR primers used for *ptp* and *ptb* cloning. Restriction sites are underlined, and regions used for Gibson cloning homologous to pET-28a(+) are highlighted in grey.

Name	Sequence (5'-3')
*ptp*F	GCTAGCATGAGCAATGAACTGCAAGAC
*ptp*R	CTCGAGTCATTTCACTATTTTTCCTATATCTCCG
*ptb*F	GCCTGGTGCCGCGCGGCAGCCATATGATGAAGAGGTTGGTGCATAAGATGAAGC
*ptb*R	GATCCGCGACCCATTTGCTGTCCACCAGTACTGCTAGCTTATTTCGATGCGGAACAAACAGCAAG
*bk*F	GCTAGCTTGCAGGAGCAGAAGTTTCGGATTTTA
*bk*R	CTCGAGTTACAAAACGGTCGCTGTCGCGAAT

7.2.2 Protein purification

Unless otherwise stated, recombinant proteins were purified by nickel-affinity chromatography as described in Section 2.7.4. Protein samples were diluted in buffers containing 5mg bovine serum albumin /ml of buffer.

7.2.2.1 Ion-exchange chromatography

Ion-exchange chromatography was used to purify the non-His-tagged BK from *E. coli* cells. It was carried out using a 5ml HiTrap™ Q HP (anion) column on an ÄKTA Explorer FPLC system. The pI value for BK was predicted to be 5.7 (ProtParam tool, ExPASy). *E. coli* BL21(DE3) cells were lysed in a 50mM citrate buffer, pH 6.2. The soluble fraction was injected onto the column at a flow rate of 1ml/min. Bound protein was eluted with the 50mM citrate buffer, pH 6.2, with a NaCl gradient of 0-1M over 120min.

7.2.3 Enzyme characterisation

7.2.3.1 Substrate specificity

The substrate specificity of the PTACs was determined by measuring activity with 50mM phosphate and 0.1mM acyl-CoA. The acyl-CoAs tested were acetyl-CoA, propionyl-CoA and butyryl-CoA, prepared using the method described in Section 2.7.7.1. The activity of each acyl-CoA was normalised to the activity with propionyl-CoA or butyryl-CoA for PTP and PTB, respectively.

The substrate specificity of BK was investigated by measuring enzyme activity in both the acid-forming and the acyl-phosphate forming directions. Enzymatic assays in the acid-forming direction were carried out as described in Section 2.7.7.4. BK activity in the acyl-phosphate forming direction was measured by a coupled assay system involving pyruvate kinase and lactate dehydrogenase (Figure 7.1). The reaction mixture (1ml) contained 50mM Hepes, pH 6.5 at 60°C, 5mM $MgCl_2.6H_2O$, 5mM KCl, 1mM ATP, 1mM phosphoenolpyruvate (PEP), 0.2mM NADH, 4.3U pyruvate kinase, 6U lactate dehydrogenase, and 1mM sodium butyrate or 1mM sodium acetate. The reaction was started by adding cell extract/enzyme, and the rate of NADH loss was measured at 340nm (ε_{340} = 6,200 $M^{-1}.cm^{-1}$).

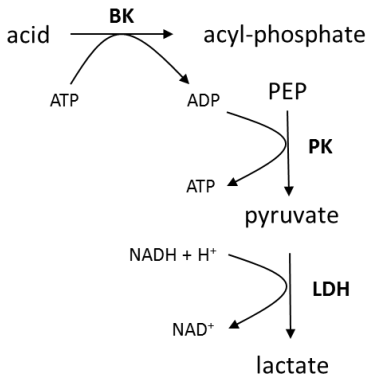


Figure 7.1 Schematic representation of the butyrate kinase coupled assay in the acyl-phosphate forming direction. The ATP production by butyrate kinase (BK) was coupled to the activities of pyruvate kinase (PK) and lactate dehydrogenase (LDH), and the rate of NADH loss was measured at 340nm.

7.2.3.2 Kinetic properties

Kinetic properties were determined for PTP and PTB by the method described in Section 2.7.7.3 and Section 3.3.1.2, respectively.

7.2.4 Creation of *ptp* and *ptb* knockout constructs

7.2.4.1 Creation of the *ptp* knockout construct

Initial attempts to create a *ptp* knockout construct were problematic. The *ptp* gene and approximately 800bp sequences flanking either side of the gene were amplified by PCR from *G. thermoglucosidasius* TM242 genomic DNA. However, it was thought that secondary structure was forming in this amplicon and this prevented the ligation of the knockout construct. Therefore, the strategy was redesigned to amplify the *ptp* gene and 660-680bp sequences flanking either side of the gene using primers ptpKOF and ptpKOR (Table 7.2). The PCR product was purified and digested with *Kpn*I and *Xba*I, prior to ligation into pUC19 which had been linearized with *Kpn*I and *Xba*I. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. Divergent primers, ptpKOdivF and ptpKOdivR (Table 7.2), were used to truncate the *ptp* gene while amplifying the pUC19 backbone and flanking regions. Purified PCR product was digested with *Acs*I, and circularised by self-ligation. Ligated plasmids were replicated by transforming *E. coli* JM109 cells.

The pUC19 backbone now contained a 1,634bp insert that consisted of the DNA sequence that flanks the *ptp* gene and a truncated version of the gene itself. This was excised from

pUC19 by digestion with *Kpn*I and *Xba*I, and ligated into pUB31 that had been linearized with the same restriction enzymes. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies that had grown on LB plates with carbenicillin.

7.2.4.2 Creation of the *ptb* knockout construct

The *ptb* gene and approximately 800bp sequences flanking either side of the gene were amplified by PCR from *G. thermoglucosidasius* TM242 genomic DNA, using primers ptbKOF and ptbKOR (Table 7.2). The PCR product was purified and digested with *Kpn*I and *Xba*I, prior to ligation into pUC19 that had been linearized with *Kpn*I and *Xba*I. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. Divergent primers, ptbKOdivF and ptbKOdivR (Table 7.2), were used to remove the *ptb* gene while amplifying the pUC19 backbone and 800bp flanking regions. Purified PCR product was digested with *Not*I, and circularised by self-ligation. Ligated plasmids were replicated by transforming *E. coli* JM109 cells.

The pUC19 backbone now contained a 1,608bp insert consisting of the DNA sequence that flanks the *ptb* gene in the *G. thermoglucosidasius* TM242 genome. This was excised from pUC19 by digestion with *Kpn*I and *Xba*I, and ligated into pUB31 that had been linearized with the same restriction enzymes. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies that had grown on LB plates with carbenicillin.

Table 7.2 PCR primers used for creation of *ptp* and *ptb* knockout constructs. Restriction sites are underlined.

Name	Sequence (5'-3')
ptpKOF	GGGGTACCCCAATCAGAATGAAGAAGGG
ptpKOR	GCTCTAGAATATGTCTTTTAACGATAGAGGCC
ptpKOdivF	GGCGCGCCGTTGGCAAAGAAAGAAAAACG
ptpKOdivR	GGCGCGCCGCCGAAAAGACTTTCTAAGTCTTC
ptbKOF	GGGGTACCGCAAATAATGGAAGTATTTTCCTTGATG
ptbKOR	GCTCTAGACCGGCAATGACCTTCGC
ptbKOdivF	AAGGAAAAAAGCGGCCGCTAAAAGAATGTTGAGGAGGAAGTC
ptbKOdivR	TTTTCCTTTTGCGGCCGCGGAATTTGTGGAGGCACTTT

7.2.5 Isolation of mutant strains with gene deletion

Mutant strains with deletion to the *ptp* and/or *ptb* genes were isolated following the protocol described in Section 2.6.16.

7.2.6 Enzyme assays

Geobacillus cells were grown in micro-aerobic tube cultures on USM medium containing 2% (w/v) glucose. Cell lysates were prepared as described in Section 2.7.3 and enzyme assays were performed as described in Section 2.7.7.3.

7.2.7 Tube fermentations

Micro-aerobic tube fermentation cultures were carried out as described in Section 2.5.1. *Geobacillus* strains were revived from glycerol stocks and grown in 2SPYNG to use as a seed culture. USM medium containing 2% (w/v) yeast extract and either 2% (w/v) glucose or 2% (w/v) xylose was inoculated with the corresponding strain once an OD_{600} of 3 had been reached in the seed culture. The cell pellet was harvested, and the supernatant analysed by HPLC, after the culture had been incubated for 48h (or 72h) at 60°C.

7.2.8 Bench-top fermentations

Batch fermentations of *Geobacillus* strains were grown on USM medium containing 2% (w/v) yeast extract and either 4% (w/v) glucose or 3% (w/v) xylose as described in Section

2.5.2. Fermentations and subsequent HPLC analysis were performed at TMO Renewables Ltd in at least duplicate runs for each strain unless otherwise stated.

7.2.9 Voges-Proskauer test

The Voges-Proskauer test was used to determine the presence of acetoin from *G. thermoglucosidasius* cultures. Cultures were grown in 2SPYNG overnight in a shaking incubator at 60°C. The cultures were centrifuged and to 1ml of supernatant 0.6ml of 5% (w/v) α-napthol and 0.4ml of 40% (w/v) KOH were added. The assay was allowed to develop for 5min, and any colour change was observed.

7.3 Results

7.3.1 Cloning

7.3.1.1 Cloning of *ptp*

The phosphotranspropionylase (PTP) encoding gene annotated by ERGO™ was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA (Figure 7.2). The 636bp DNA fragment was excised from the gel, cloned into the pGEM®-T vector and transformed into *E. coli* JM109 electrocompetent cells as described in Section 7.2.1.1. The gene was sub-cloned into the pET-28a(+) expression vector using restriction enzymes *Nhe*I-HF and *Xho*I. Plasmid DNA was purified, and sequencing confirmed that the correct nucleotide sequence was present.

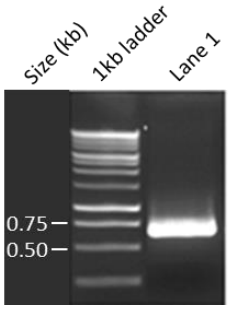


Figure 7.2 Agarose gel electrophoresis of the *ptp* amplicon. The 636bp PCR-amplified gene (lane 1) was visualised alongside the 1kb ladder (NEB, Massachusetts, US).

7.3.1.2 Cloning of *ptb*

The phosphotransbutyrylase (PTB) encoding gene annotated by ERGO™ was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA (Figure 7.3). The 921bp DNA fragment was cloned using the Gibson isothermal enzymic DNA assembly method (Gibson *et al.* 2009). The resulting *ptb*-pET-28a(+) construct was purified, and sequencing confirmed that the correct nucleotide sequence was present.

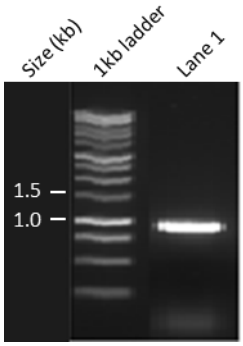


Figure 7.3 Agarose gel electrophoresis of the *ptb* amplicon. The 921bp PCR-amplified gene (lane 1) was visualised alongside the 1kb ladder (NEB, Massachusetts, US).

7.3.1.3 Cloning of *bk*

The butyrate kinase (BK) encoding gene annotated by ERGO™ was PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA (Figure 7.4). The 1,110bp DNA fragment was excised from the gel, cloned into the pGEM®-T vector and transformed into *E. coli* JM109 electrocompetent cells as described in Section 7.2.1.3. The gene was sub-cloned into the pET-28a(+) expression vector using restriction enzymes *Nhe*I and *Xho*I. Plasmid DNA was purified, and sequencing confirmed that the correct nucleotide sequence was present. Due to precipitation of the recombinant protein (see Section 7.3.2) the gene was also sub-cloned into the pET-24a(+) expression vector using restriction enzymes *Nhe*I and *Xho*I. Expression from the pET-24a(+) construct would produce a recombinant protein without a His-tag, unlike expression from pET-28a(+).

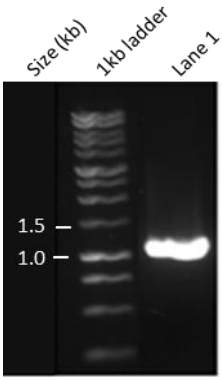


Figure 7.4 Agarose gel electrophoresis of the *bk* amplicon. The 1,110bp PCR-amplified gene (lane 1) was visualised alongside the 1kb ladder (NEB, Massachusetts, US).

7.3.2 Expression and purification of recombinant enzymes

The putative *ptp*, *ptb* and *bk* genes were expressed in *E. coli* BL21(DE3) cells using the inducible T7 expression system. This resulted in recombinant N-terminal His-tagged proteins with molecular masses close to the predicted molecular masses of 25.9kDa, 34.7kDa and 42.3kDa for PTP, PTB, and BK, respectively. Proteins were purified to homogeneity by nickel-affinity chromatography and visualised by SDS-PAGE (Figure 7.5).

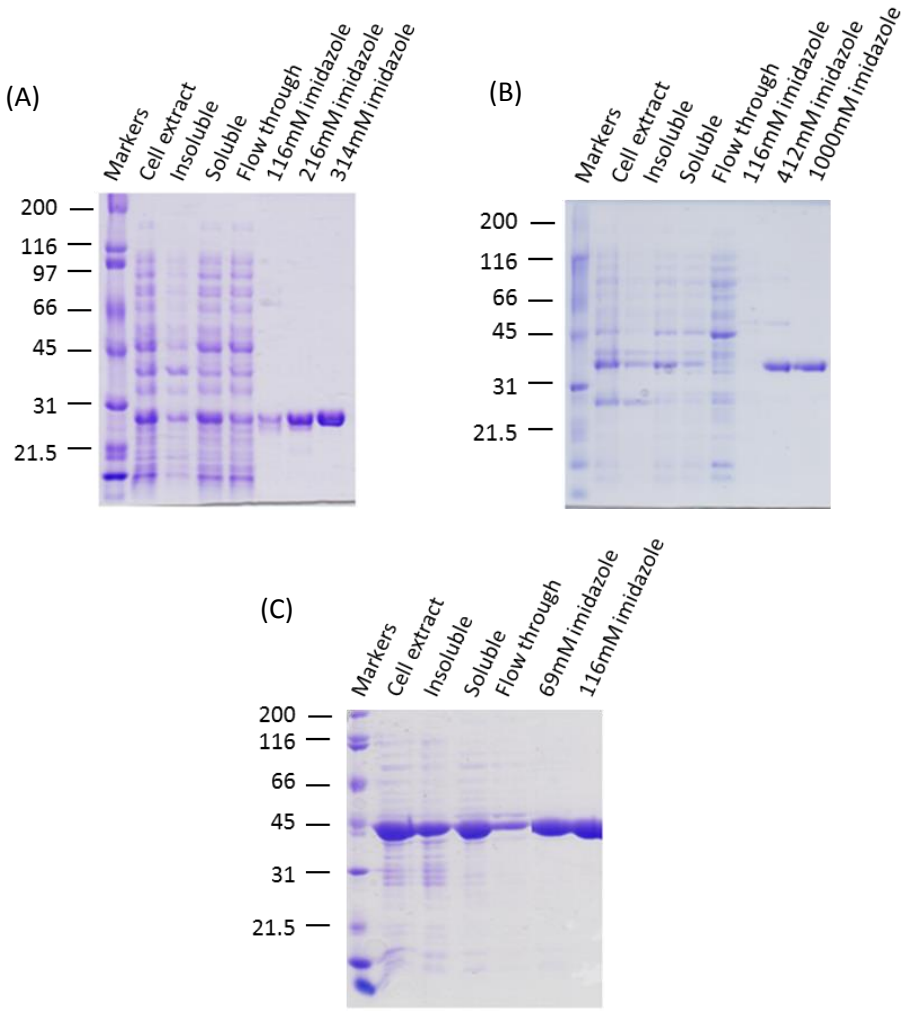


Figure 7.5 SDS-PAGE analysis of recombinant His-tagged proteins. His-tagged PTP (A), PTB (B), and BK (C) with expected molecular masses of 25.9kDa, 34.7kDa and 42.3kDa, respectively, were purified from *E. coli* BL21(DE3) cell lysates by nickel affinity chromatography. The soluble fractions of cell extracts were separated from the insoluble debris by centrifugation. Soluble samples were loaded onto the column and the proteins were purified by washes with increasing imidazole concentrations. The size of the marker proteins are in kDa.

During the nickel-affinity chromatography purification of PTB and BK, the His-tagged proteins precipitated. Therefore, although the proteins were purified to homogeneity as observed by SDS-PAGE, a significant amount of protein was lost during the purification. Further precipitation was also observed when the purified proteins were subjected to buffer exchange by dialysis in order to remove the imidazole. The non-His-tagged recombinant BK protein expressed from the pET-24a(+) construct did not precipitate during anion-exchange chromatography purification (Figure 7.6). This suggests that precipitation of the His-tagged BK is a result of either imidazole or the effects of introducing a poly(His) N-terminal peptide. Despite preventing precipitation, however, anion-exchange chromatography did not purify the BK to homogeneity. Therefore further characterisation was conducted on purified His-tagged protein samples.

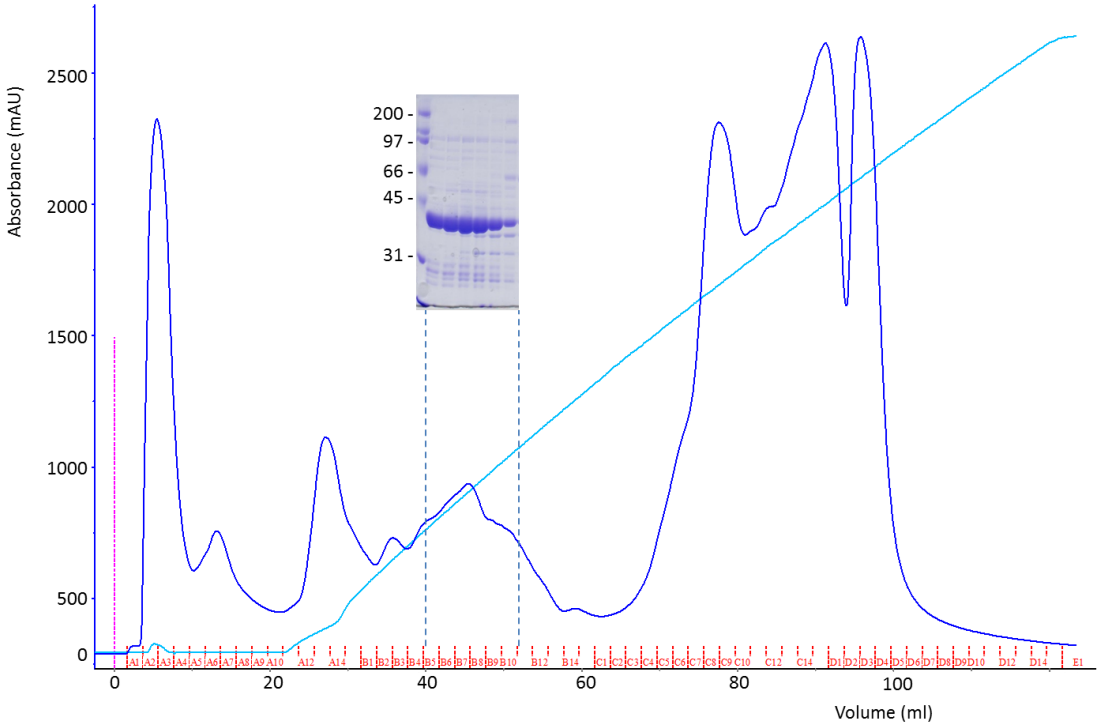


Figure 7.6 Anion exchange chromatograph for the purification of BK from *E. coli* cell lysate. Non-His tagged BK, with an expected molecular mass of 39.9kDa was purified by anion exchange chromatography. The dark blue line is the absorbance at 280nm (mAU), the light blue line represents the increase in NaCl concentration, and the fractions where the majority of the BK eluted are indicated with the dashed lines and the SDS-PAGE gel. The size of the marker proteins are in kDa.

7.3.3 Characterisation of the recombinant phosphotransacylases

7.3.3.1 Substrate specificity

Both PTP and PTB had a degree of promiscuity to their activity with acyl-CoAs. Both recombinant PTACs could catalyse the acyl-group transfer to orthophosphate with acetyl-CoA, propionyl-CoA and butyryl-CoA (Figure 7.7). The PTP could catalyse the transfer with acetyl-CoA at the same rate of catalysis with propionyl-CoA. However, the PTB could only catalyse the transfer with acetyl-CoA at 20% of the rate of catalysis with butyryl-CoA. It is worth comparing the activity of PTP and PTB with that of PTA (Figure 4.22), which could catalyse the reaction with acetyl-CoA at the same rate as with propionyl-CoA, but could only use butyryl-CoA at 0.1% of the rate with acetyl-CoA.

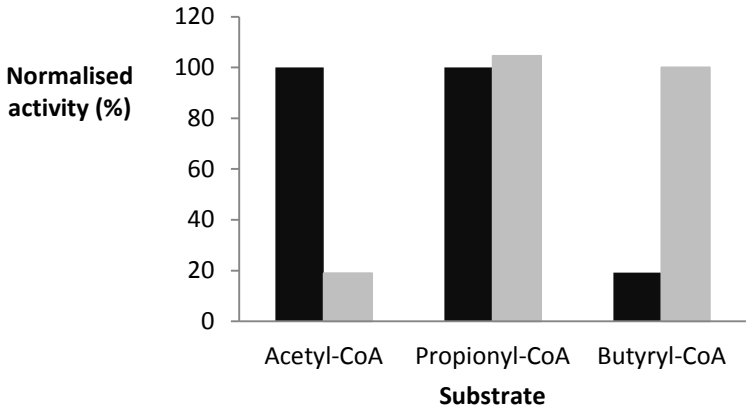


Figure 7.7 Substrate specificity of PTP and PTB. Purified recombinant PTP (black bars) and PTB (grey bars) were assayed with 50mM phosphate and either 0.1mM acetyl-CoA, propionyl-CoA, or butyryl-CoA. Activity was normalised to 100% with propionyl-CoA for PTP and butyryl-CoA for PTB.

7.3.3.2 Kinetic properties with annotated substrate

The kinetic characterisation of the PTACs was performed using the substrates orthophosphate and propionyl-CoA or butyryl-CoA for PTP and PTB, respectively. Both purified recombinant proteins were diluted in a 50mM Tris buffer, pH 8.0, containing 5mg BSA/ml; the PTB dilution buffer also contained 412mM imidazole. The PTP appeared to display positive cooperative binding towards orthophosphate concentration (Figure 7.8 A and B); data were fitted to the Hill equation and yielded a $V_{\max}$ of 272 (± 12) U/mg, a $K_{0.5(\text{phosphate})}$ of 27 (± 2) mM, and an n of 1.7 (where n is the Hill coefficient). PTP displayed typical Michaelis-Menten kinetics towards propionyl-CoA concentration (Figure 7.8 C and D); data were analysed by SigmaPlot and yielded a $V_{\max}$ of 242 (± 6) U/mg and a $K_{\text{m(propionyl-CoA)}}$ of 0.11 (± 0.01) mM.

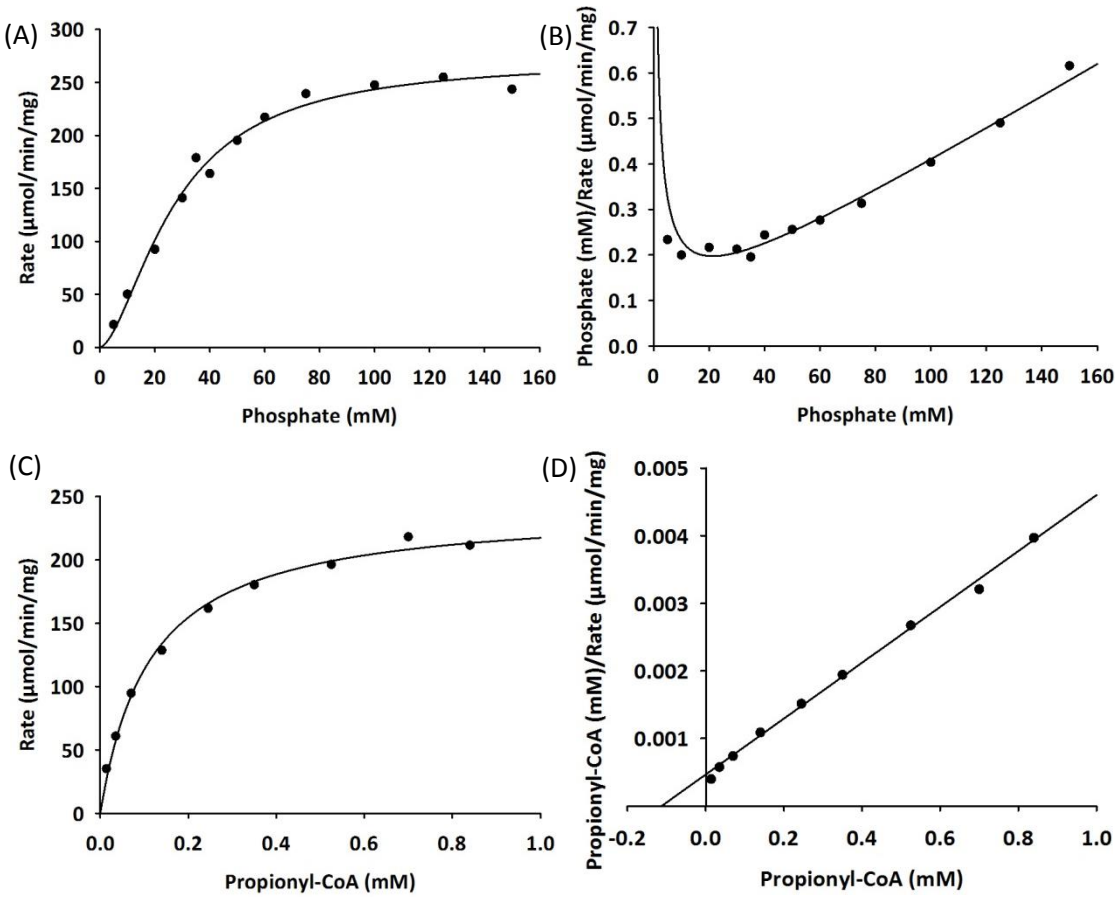


Figure 7.8 Phosphotranspropionylase activity in the propionyl-CoA forming direction. The propionyl-CoA was kept constant at 0.42mM while the phosphate was varied; data are presented as (A) a Michaelis-Menten plot and (B) a Hanes-Woolf plot. The phosphate was kept constant at 50mM while the propionyl-CoA was varied; data are presented as (C) a Michaelis-Menten plot and (D) a Hanes-Woolf plot.

To determine the kinetic parameters for PTB, a matrices approach was taken as described in Section 3.3.1.2. The PTB displayed typical Michaelis-Menten kinetics towards orthophosphate concentration at each fixed concentration of butyryl-CoA (Figure 7.9 A). Secondary plots were used to determine a $V_{\max}$ of 294 (± 10) U/mg and a $K_{\mathrm{m(butyryl\text{-}CoA)}}$ of 0.0077 (± 0.0013) mM (Figure 7.9 B and C). PTB also displayed typical Michaelis-Menten kinetics towards butyryl-CoA concentration at each fixed concentration of orthophosphate (Figure 7.10 A). Secondary plots were used to determine a $V_{\max}$ of 292 (± 24) U/mg and a $K_{\mathrm{m(phosphate)}}$ of 4.4 (± 1.3) mM (Figure 7.10 B and C).

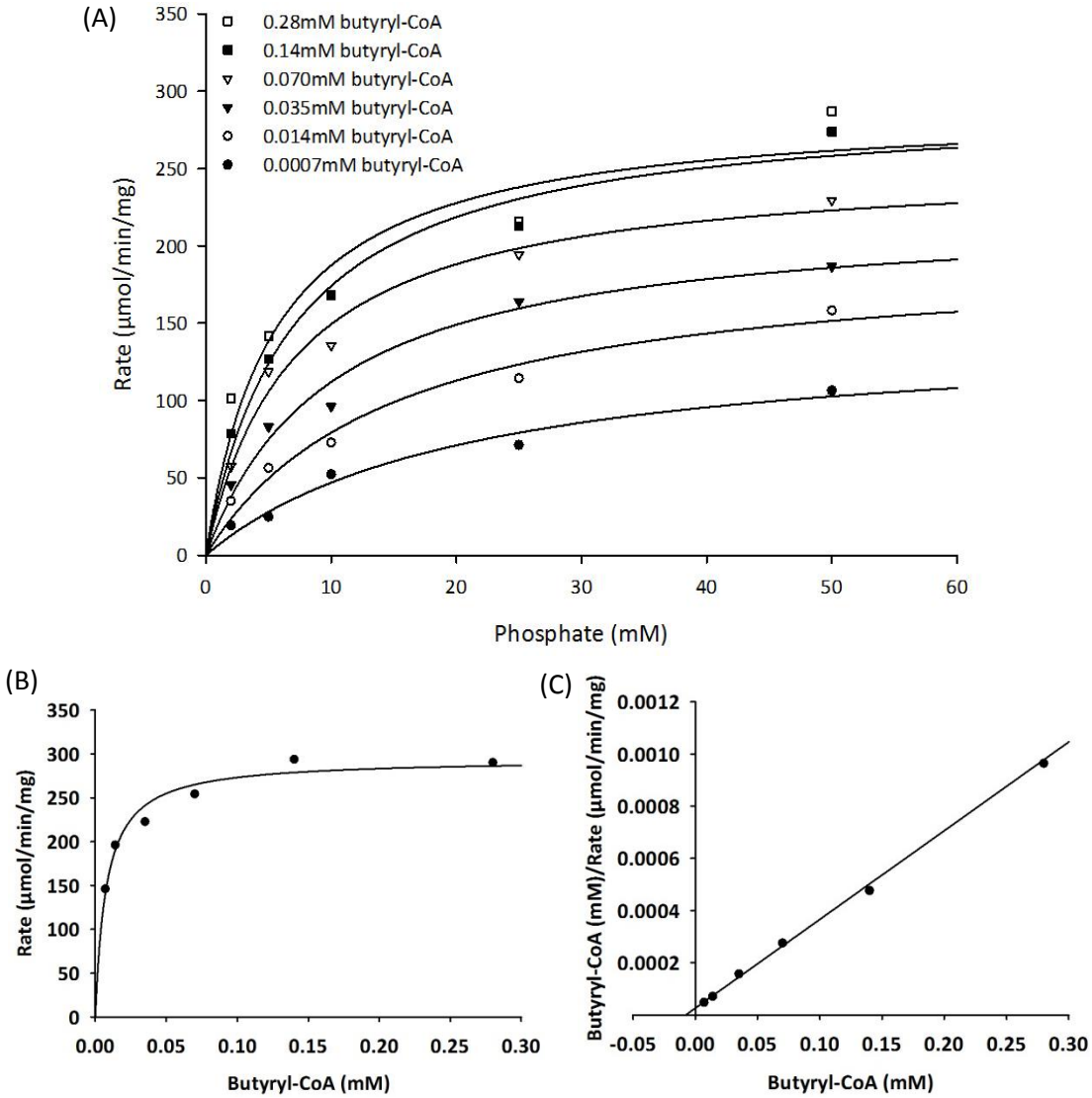


Figure 7.9 Phosphotransbutyrylase activity in the butyryl-CoA forming direction. The relationship between specific activity (U/mg) and phosphate concentration at a variety of fixed butyryl-CoA concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true $\boldsymbol{K_m}$ for butyryl-CoA at saturating phosphate concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

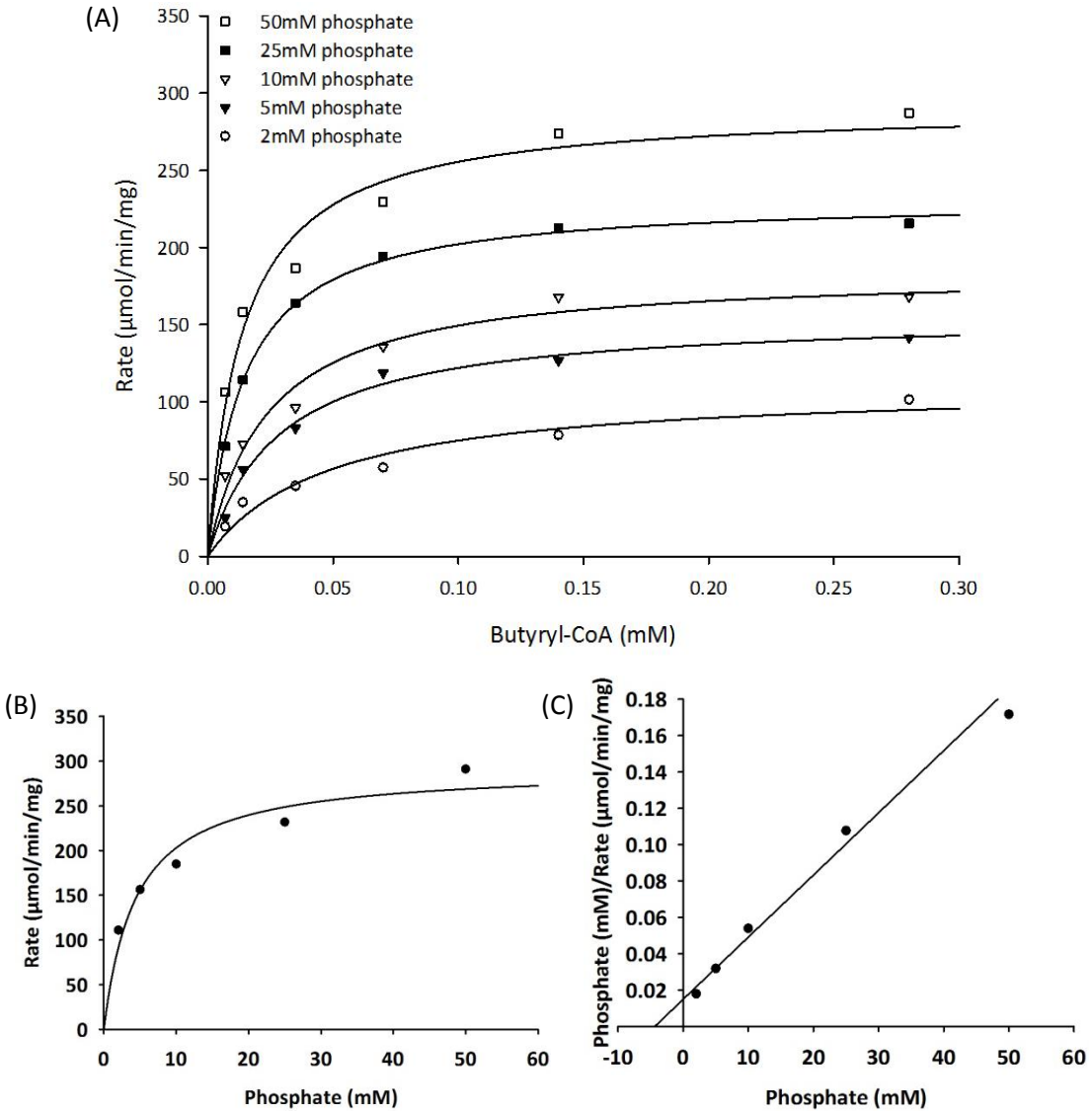


Figure 7.10 Phosphotransbutyrylase activity in the butyryl-CoA forming direction. The relationship between specific activity (U/mg) and butyryl-CoA concentration at a variety of fixed phosphate concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true $\boldsymbol{K_m}$ for phosphate at saturating butyryl-CoA concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

7.3.3.3 Kinetic properties with acetyl-CoA

Both PTP and PTB were able to catalyse the conversion of acetyl-CoA and orthophosphate to acetyl-phosphate and CoA. The kinetic properties of the PTACs were determined with orthophosphate and acetyl-CoA. The PTP appeared to display positive cooperative binding towards orthophosphate concentration (Figure 7.11 A and B); data were fitted to the Hill equation and yielded a $V_{\max}$ of 250 (± 19) U/mg, a $K_{0.5(\text{phosphate})}$ of 21 (± 3) mM, and an n of 2.6. PTP displayed typical Michaelis-Menten kinetics towards acetyl-CoA concentration (Figure 7.11 C and D); data were analysed by SigmaPlot and yielded a $V_{\max}$ of 269 (± 11) U/mg and a $K_{\text{m(acetyl-CoA)}}$ of 0.17 (± 0.02) mM.

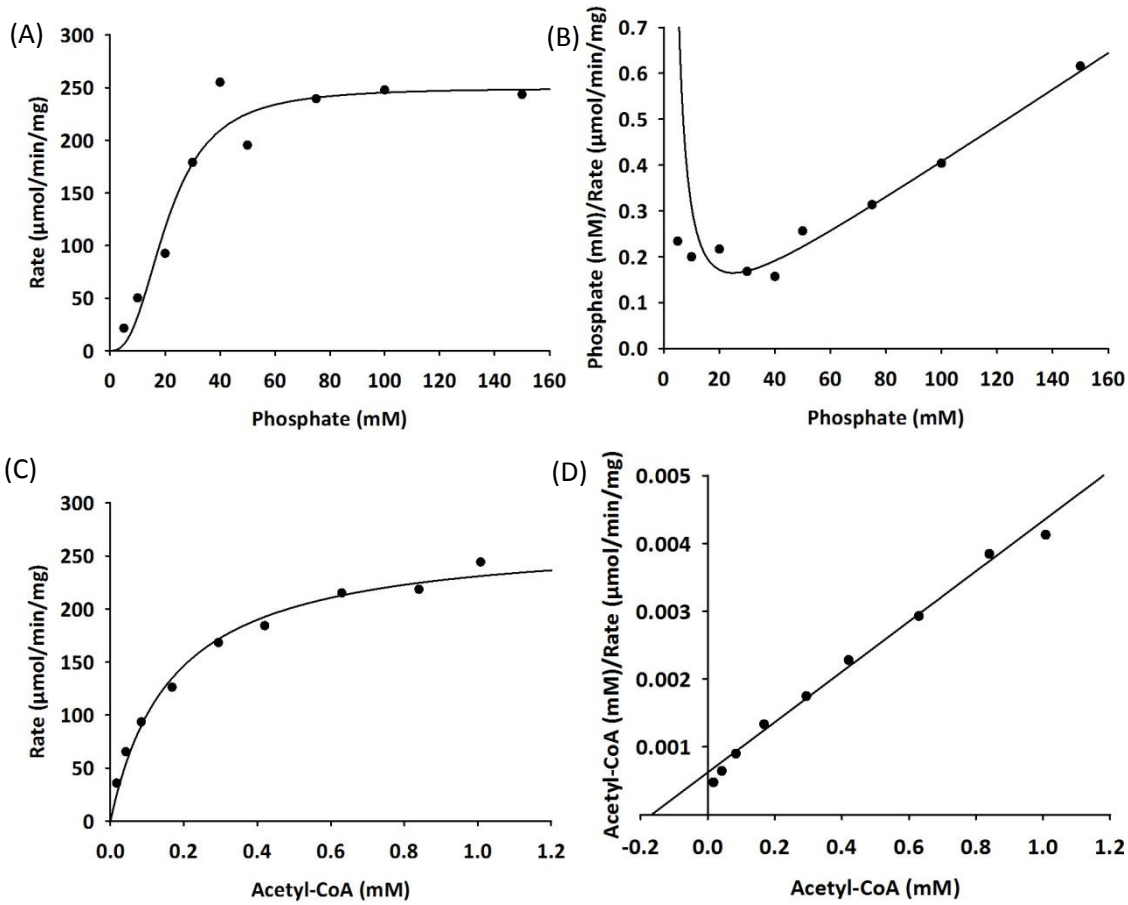


Figure 7.11 Phosphotranspropionylase activity in the acetyl-CoA forming direction. The acetyl-CoA was kept constant at 0.50mM while the phosphate was varied; data are presented as (A) a Michaelis-Menten plot and (B) a Hanes-Woolf plot. The phosphate was kept constant at 50mM while the acetyl-CoA was varied; data are presented as (C) a Michaelis-Menten plot and (D) a Hanes-Woolf plot.

To determine the kinetic parameters for PTB a matrices approach was taken as described in Section 3.3.1.2. The PTB displayed typical Michaelis-Menten kinetics towards acetyl-CoA concentration at each fixed concentration of orthophosphate (Figure 7.12 A). Secondary plots were used to determine a $V_{\max}$ of 260 (± 66) U/mg and a $K_{\mathrm{m(acetyl\text{-}CoA)}}$ of 0.067 (± 0.046) mM (Figure 7.12 B and C). PTB also displayed typical Michaelis-Menten kinetics towards orthophosphate concentration at each fixed concentration of acetyl-CoA (Figure 7.13 A). Secondary plots were used to determine a $V_{\max}$ of 319 (± 19) U/mg and a $K_{\mathrm{m(phosphate)}}$ of 20 (± 3) mM (Figure 7.13 B and C).

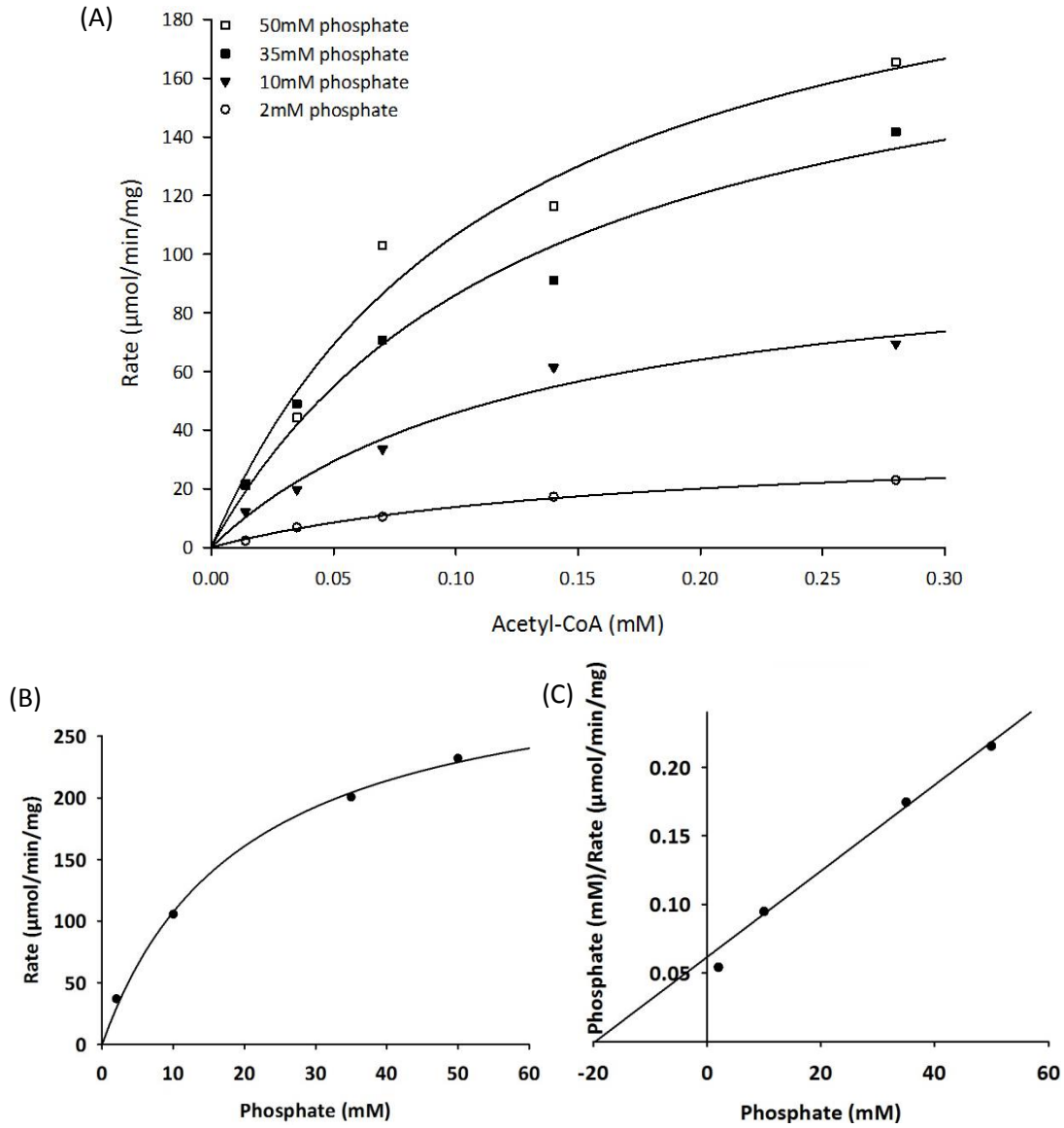


Figure 7.12 Phosphotransbutyrylase activity in the acetyl-CoA forming direction. The relationship between specific activity (U/mg) and acetyl-CoA concentration at a variety of fixed phosphate concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true K_m for phosphate at saturating acetyl-CoA concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

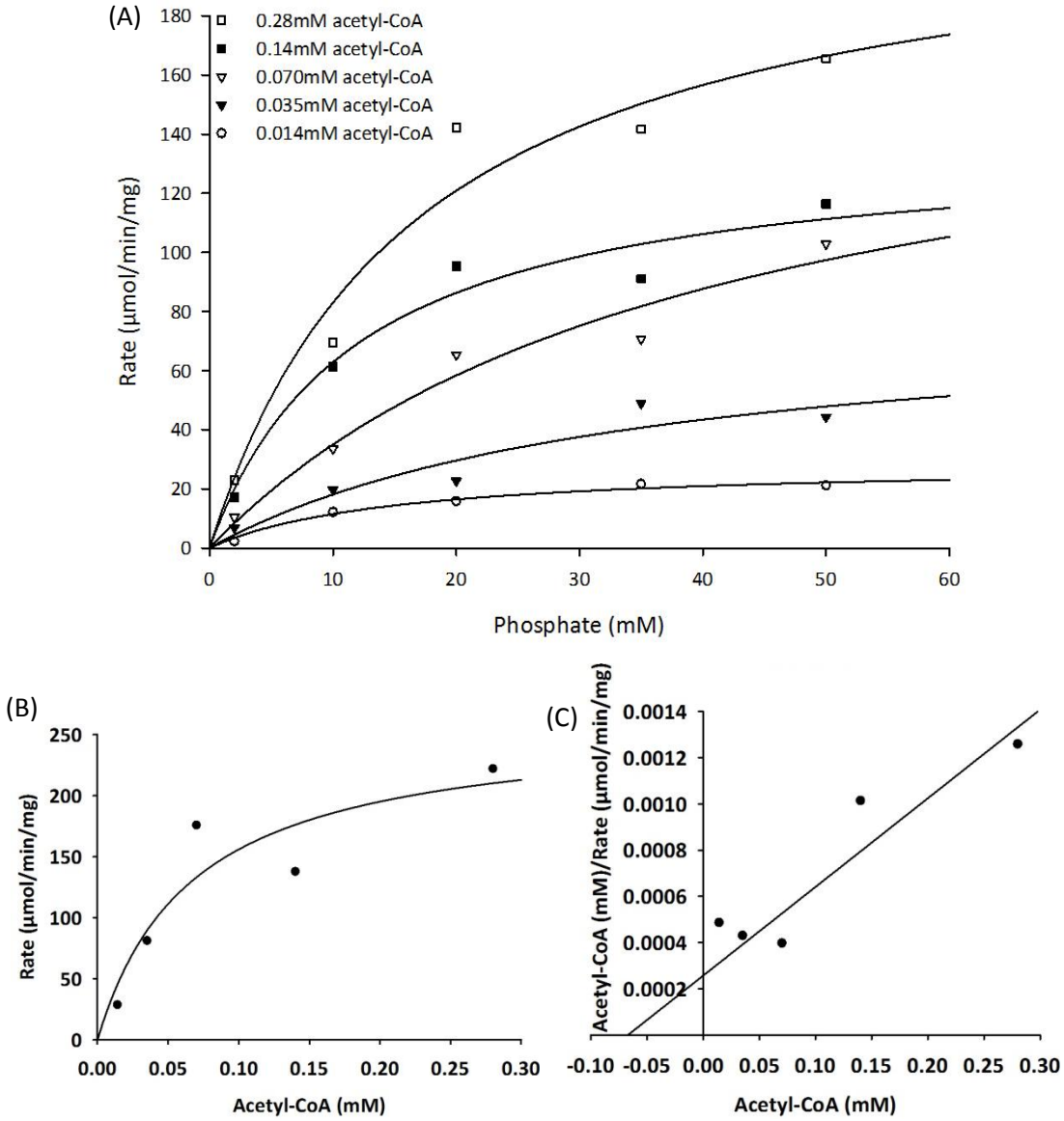


Figure 7.13 Phosphotransbutyrylase activity in the acetyl-CoA forming direction. The relationship between specific activity (U/mg) and phosphate concentration at a variety of fixed acetyl-CoA concentrations is displayed as (A) a Michaelis-Menten plot. The secondary plots were used to determine a true $\boldsymbol{K_m}$ for acetyl-CoA at saturating phosphate concentration in the form of (B) a Michaelis-Menten plot and (C) a Hanes-Woolf plot.

7.3.4 Substrate specificity of the recombinant butyrate kinase

The observation that PTB could use acetyl-CoA in addition to butyryl-CoA as a substrate led to the hypothesis that butyrate kinase (BK) could accept acetyl-phosphate in addition to butyryl-phosphate. At the time of writing, butyryl-phosphate was not commercially available; therefore, the BK was measured in the reverse direction with butyrate and ATP as described in Section 7.2.3.1. The rate of catalysis with butyrate was directly proportional to the volume of BK added to the enzyme assay (Figure 7.14), indicating that initial rates were being measured. There was, however, no activity detected with acetate and ATP as

substrates. A separate aliquot of recombinant BK was assayed in the acetate-forming direction and activity was detected with acetyl-phosphate and ADP. However, this sample was not purified to homogeneity, and therefore the presence of an *E. coli* acetate kinase cannot be ruled out. The substrate specificity of BK was not pursued any further in this Thesis.

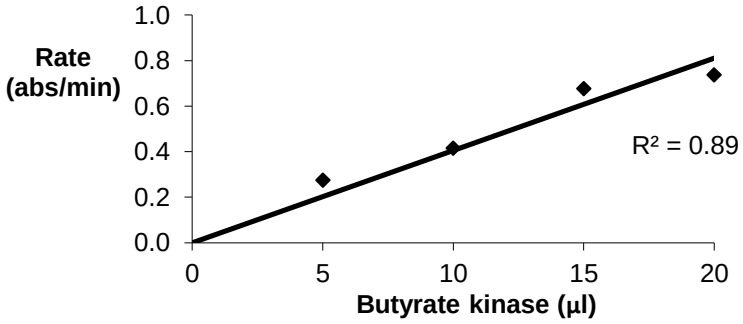


Figure 7.14 The relationship between rate of catalysis and volume of BK in the assay. BK was assayed in the butyryl-phosphate forming direction as described in the Methods.

7.3.5 Deletion of *ptp* and *ptb*, and strain characterisation

7.3.5.1 Creation of the knockout strains

The pUB31-based constructs designed to delete the *ptp* and *ptb* genes from *G. thermoglucosidasius* were created in three stages, following a similar strategy to that used to create the *pta*KO-pUB31 construct, as described in Section 6.3.2.1. The procedure outlined in Figure 6.3 was used to delete *ptp* and *ptb* from *G. thermoglucosidasius* TM242, resulting in the Δ*ptp* and Δ*ptb* strains, respectively. The *ptp* and *ptb* genes were also deleted from the Δ*pta* background strain, creating the double PTAC knockouts Δ*pta*Δ*ptp* and Δ*pta*Δ*ptb*. Finally the *ptp* gene was deleted from the Δ*ptb* and Δ*pta*Δ*ptb* background strains to create the double PTAC knockout Δ*ptp*Δ*ptb* and the triple PTAC knockout Δ*pta*Δ*ptp*Δ*ptb*, respectively. Colony PCR was used to identify successful gene knockouts (Figure 7.15). It is worth noting that the PTAC gene knockouts are all in the *G. thermoglucosidasius* TM242 genetic background.

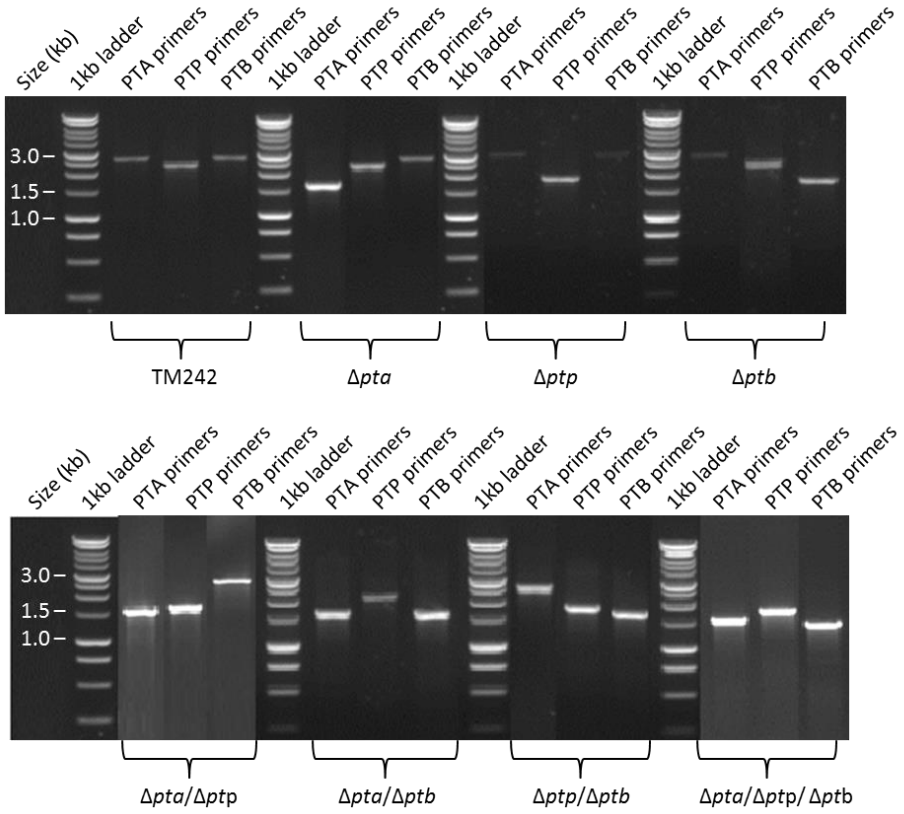


Figure 7.15 PCR colony screen to identify kanamycin sensitive double crossover mutants. Kanamycin sensitive colonies were screened by PCR with primers that anneal to the regions of DNA flanking the *pta* (PTA primers), *ptp* (PTP primers), and *ptb* (PTB primers). The TM242 strain has intact genes for *pta*, *ptp*, and *ptb* evident by amplicons of approximately 2.7kb, 2.3kb and 2.7kb, respectively. Strains with successful deletions to *pta* or *ptb* are evident by an amplicon of approximately 1.6kb, while strains with successful deletions to *ptp* are evident by an amplicon of approximately 1.8kb.

7.3.5.2 Phosphotransacylase activities

Enzymic activities of the TM242 strain and the seven different combinations of PTAC knockouts were analysed from micro-aerobically grown cells (Figure 7.16). The Δ*pta* strain had 5% activity with acetyl-CoA compared to the TM242 strain (Figure 7.16 A). The double PTAC knockout Δ*pta*Δ*ptp* also had around 5% activity with acetyl-CoA compared to the TM242 strain. No activity was observed with acetyl-CoA from cell lysates of the double PTAC knockout Δ*pta*Δ*ptb*, or the triple PTAC knockout (Δ*pta*Δ*ptp*Δ*ptb*). These data suggest that PTP was not present at detectable levels in either the Δ*pta* or Δ*pta*Δ*ptb* strains, and that PTB is the enzyme responsible for the remaining PTAC activity with acetyl-CoA in the Δ*pta* strain.

Activity with propionyl-CoA from a Δ*ptp* cell lysate was comparable with TM242 (Figure 7.16 B), although with a Δ*pta* cell lysate 15% of the activity was detected. Moreover, the

Δ*pta*Δ*ptb* strain had no activity. These data indicate that PTP was not present at detectable levels in the TM242 strain, and that PTA is responsible for the majority of activity detected with propionyl-CoA in TM242, with PTB accounting for the activity detected in the Δ*pta* strain. An interesting observation is that the activity with both acetyl-CoA and propionyl-CoA has increased in the Δ*ptp*Δ*ptb* strain, suggesting that PTA activity and/or expression has increased in this strain when compared to TM242.

Any strain carrying a deletion to *ptb* did not catalyse the reaction between orthophosphate and butyryl-CoA (Figure 7.16 C). All strains that contain an intact *ptb* gene could catalyse the reaction with butyryl-CoA to comparable levels to TM242. This observation suggests that, under the growth conditions used, PTB is the sole enzyme responsible for catalysing the acyl-group transfer between butyryl-CoA and phosphate. Considering that PTA cannot use butyryl-CoA but PTP can, this also supports the evidence that PTP was not expressed in these strains.

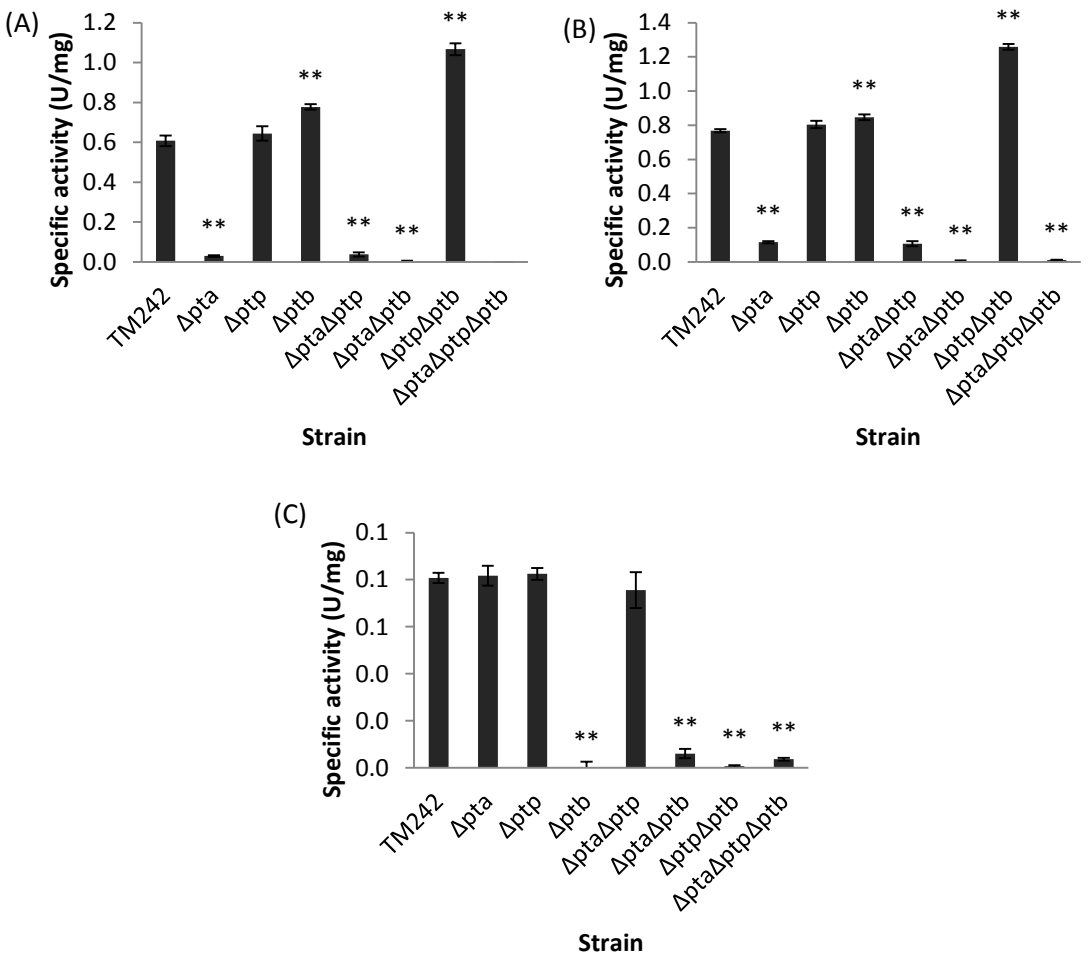


Figure 7.16 Phosphotransacylase activity in cell lysates of *G. thermoglucosidasius*. Cell lysates from micro-aerobically grown cultures were assayed with 50mM phosphate and 0.1mM of (A) acetyl-CoA, (B) propionyl-CoA or (C) butyryl-CoA. Each assay was performed in triplicate and the relevant background rate was subtracted. The student's t-test was used to determine the significance of any differences between the rate of activity from TM242 cell lysate compared to the activity from the PTAC knockout cell lysates (one star indicates a p-value <0.05, two stars indicate a p-value <0.01 at 95% confidence limits).

7.3.5.3 Aerobic growth

The ability to grow aerobically was confirmed by growing the PTAC knockout strains in baffled conical flasks (Figure 7.17). All strains were capable of growth to a healthy OD, but the lag phase was particularly affected in cultures with the Δ*ptp*Δ*ptb* strain. The deletion to *pta* resulted in increased growth rate of the culture, except to strains also carrying a deletion to *ptb*.

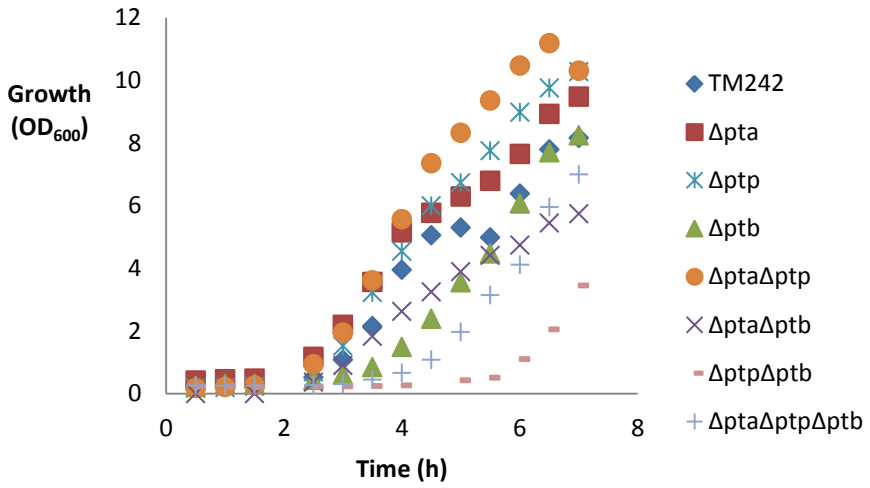


Figure 7.17 Aerobic growth of *G. thermoglucosidasius* TM242 and the PTAC knockout strains on 2SPYNG. The cultures were grown in baffled conical flasks in a shaking incubator and the absorbance at 600nm was recorded.

7.3.5.4 Strain characterisation

Ethanol and organic acid yields from micro-aerobic tube cultures of *G. thermoglucosidasius* strains on USM medium containing either glucose or xylose were used to analyse the effects of the PTAC knockouts (Figure 7.18). All strains grown on glucose had consumed all the available sugar by the end of the fermentation (48h), and all the PTAC knockout strains produced ethanol in yields comparable to TM242 (Figure 7.18 A). All the strains containing at least one knockout of a PTAC-encoding gene produced lower acetate yields under glucose growth. The triple PTAC knockout produced the lowest yield, which was 30% less than that for TM242. However, the carbon flux in this triple PTAC knockout strain seems to be diverted to lactate production, despite the *ldh* gene disruption in the TM242 background; the triple PTAC knockout strain showed a greater than 5-fold increase in lactate yield compared to the TM242 strain.

The effects of the PTAC knockouts were more pronounced when they were cultured on USM containing xylose (Figure 7.18 C). The most noticeable differences were observed with the Δ*pta*Δ*ptb* and triple PTAC knockout (Δ*pta*Δ*ptp*Δ*ptb*) strains. These two strains did not consume all the available xylose after 48h, the Δ*pta*Δ*ptb* and triple PTAC knockout strains consumed 40% and 20% of the sugar, respectively. Sugar consumption was not improved by incubating cultures for a longer period of 72h, indicating that growth of these strains was arrested as opposed to them having a longer lag phase. The ethanol yield was lower for all PTAC mutants by no more than 10%, except for the Δ*pta*Δ*ptp*, the Δ*pta*Δ*ptb* and the triple

PTAC knockout strains. These strains had ethanol yields lowered by 15%, 57% and 72% respectively, compared to TM242. Acetate yields were comparable to TM242, except for the Δ*pta*Δ*ptp* strain, which decreased by 21% and the Δ*pta*Δ*ptb* and the triple PTAC knockout strains which increased 1.6- and 2.5-fold respectively. The Δ*pta*Δ*ptb* and the triple PTAC knockout strains also produced significantly greater yields of pyruvate and lactate when compared to TM242.

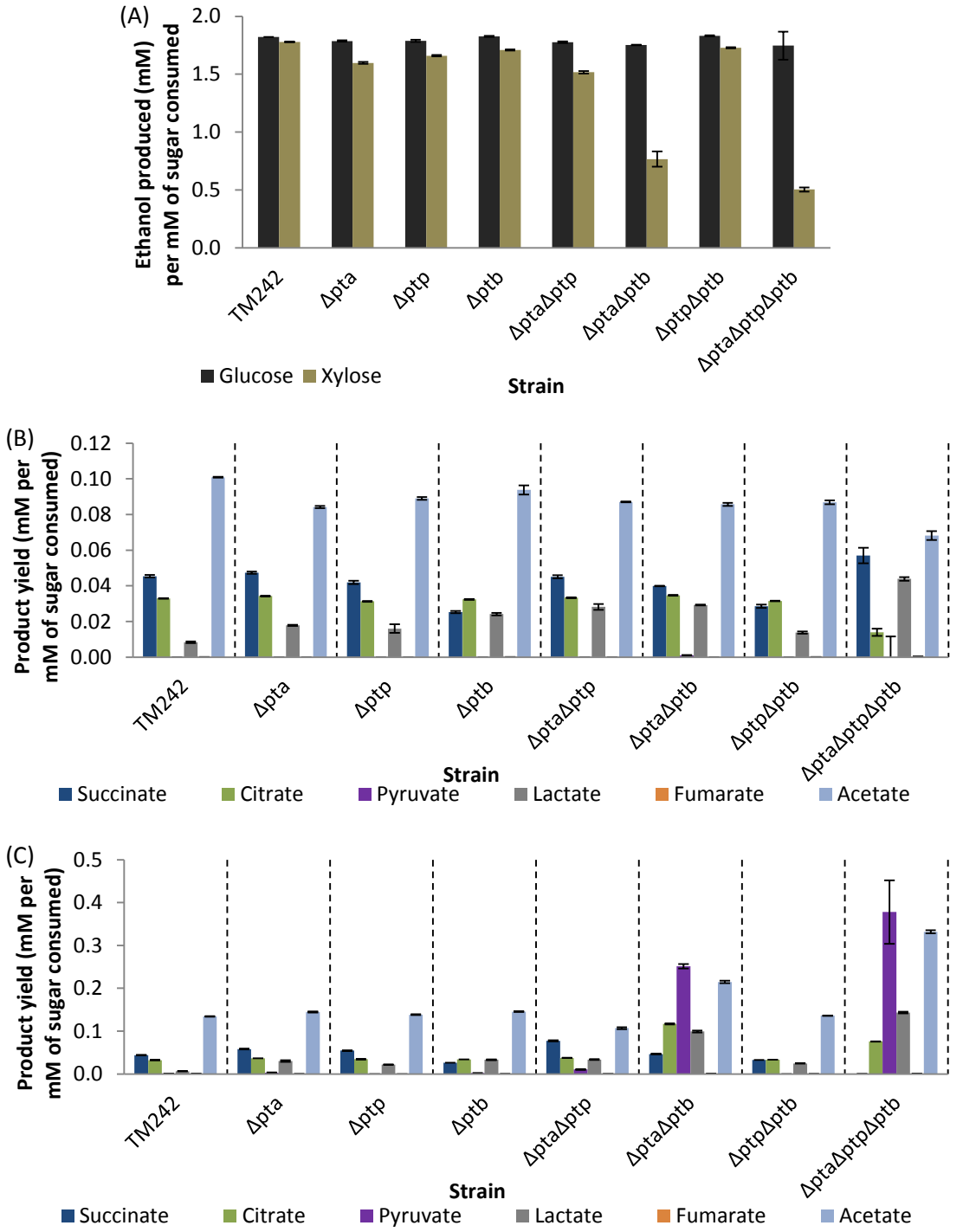


Figure 7.18 End-point products of fermentation by *G. thermoglucosidasius*. Metabolites were determined from tube fermentation cultures of TM242 and the PTAC knockout strains grown on USM containing 2% (w/v) of either glucose or xylose. Cultures were grown at 60°C with shaking for 48h in sealed falcon tubes to induce fermentative metabolism. Products are given as a yield in terms of mM of product per mM of sugar consumed. Bar charts represent the ethanol yields on glucose and xylose (A), and the organic acid profiles for glucose grown cells (B) and xylose grown cells (C).

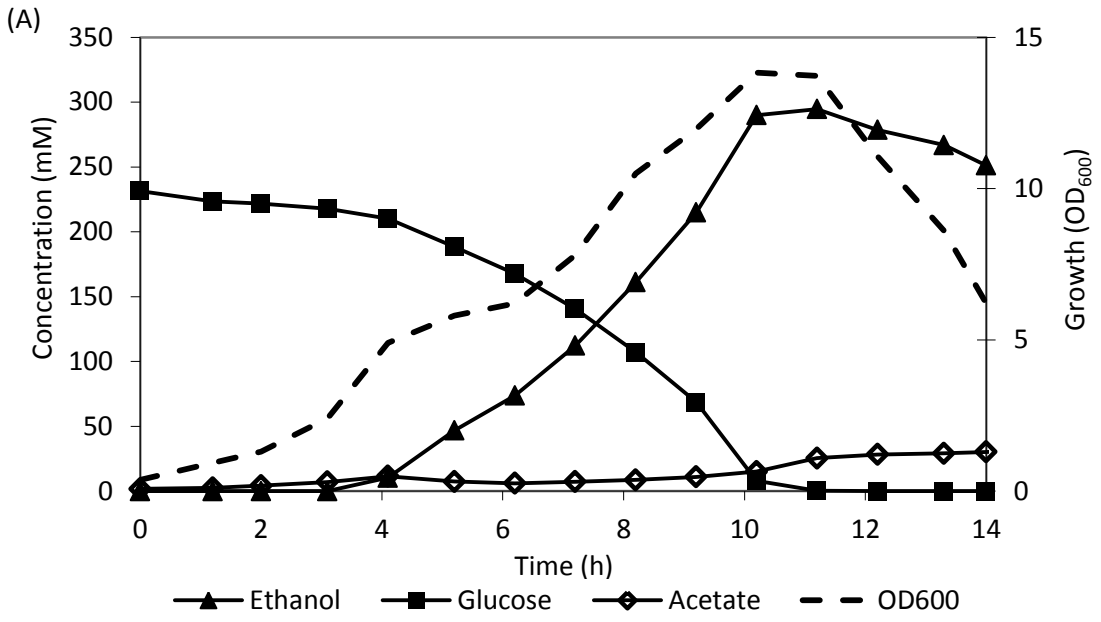
Batch fermentation profiles of the *G. thermoglucosidasius* strains grown on USM containing 4% (w/v) glucose are shown in Figure 7.19. Of the single PTAC knockout strains, the Δ*pta* strain grew the fastest and produced the highest peak ethanol concentration. Although the

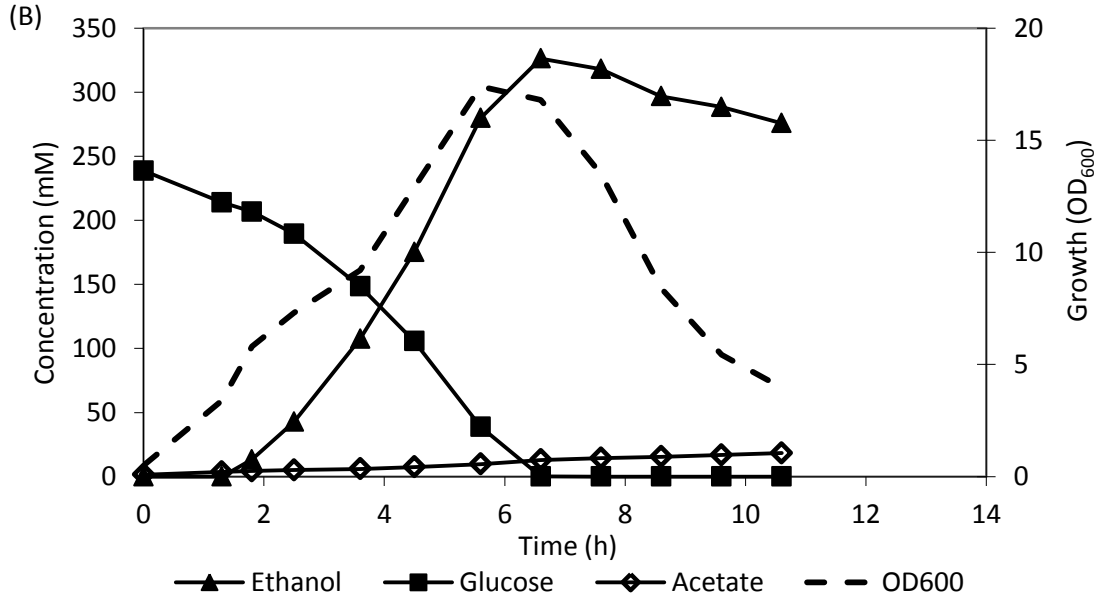
peak ethanol concentration for the Δ*ptp* strain was not recorded, the growth was reduced and more pyruvate was secreted by this strain. The Δ*ptb* strain had a similar fermentation profile to the TM242 strain.

The glucose fermentation profiles for the double PTAC knockout and triple PTAC knockout strains are shown in Figure 7.20. Unfortunately, consistent data were not obtained for the Δ*pta*Δ*ptp* strain so it is not shown, and only a single tank was run with the Δ*pta*Δ*ptb* strain. The Δ*pta*Δ*ptb* strain was a challenging strain to grow, as indicated by the slow increase in OD_{600}. Replicate data would be required to make a definitive conclusion with this strain. The Δ*ptp*Δ*ptb* strain grew rapidly, and the fermentation profile was similar to TM242. The triple PTAC knockout grew slowly and did not produce high concentrations of ethanol, although there was an increased concentration of lactate produced. Despite the deletion of the three PTACs, acetate was still produced by this strain.

Notable alterations to the acetate profile were observed with strains carrying deletions to the *pta* gene (Figure 7.21). The Δ*pta* and the Δ*pta*Δ*ptp*Δ*ptb* strains did not re-assimilate any acetate generated. In contrast the TM242 acetate profile showed an initial increase in acetate, peaking at 15mM, which subsequently dropped to around 7mM. The end-point concentration of acetate for TM242 was around 40mM, suggesting the acetate is generated again; a similar profile is also observed for the Δ*ptp* and Δ*ptb* strains. These results indicate that a strain carrying a deletion to *pta* still produces acetate but cannot re-assimilate any acetate produced.

G. thermoglucosidasius strains carrying deletions to *pta* have reduced activity with acetyl-CoA and orthophosphate, but still generate acetate, suggesting an alternative acetate-producing pathway exists. It has been reported that the Gram-positive mesophile, *B. subtilis*, generates acetate via acetoin and butanediol (Speck and Freese 1973). The Voges-Proskauer test was used on supernatants from *G. thermoglucosidasius* cultures, and indicated that acetoin accumulation increased from cultures of Δ*pta* and Δ*pta*Δ*ptb* (Figure 7.22). Results with cultures of TM242 and Δ*ptb* were negative indicating an absence of acetoin, and suggesting that the disruption of *pta* results in an increased flux from pyruvate to acetoin.


(A)
Concentration (mM)
Growth (OD600)
Time (h)
Ethanol
Glucose
Acetate
OD600


(B)
Concentration (mM)
Growth (OD600)
Time (h)
Ethanol
Glucose
Acetate
OD600

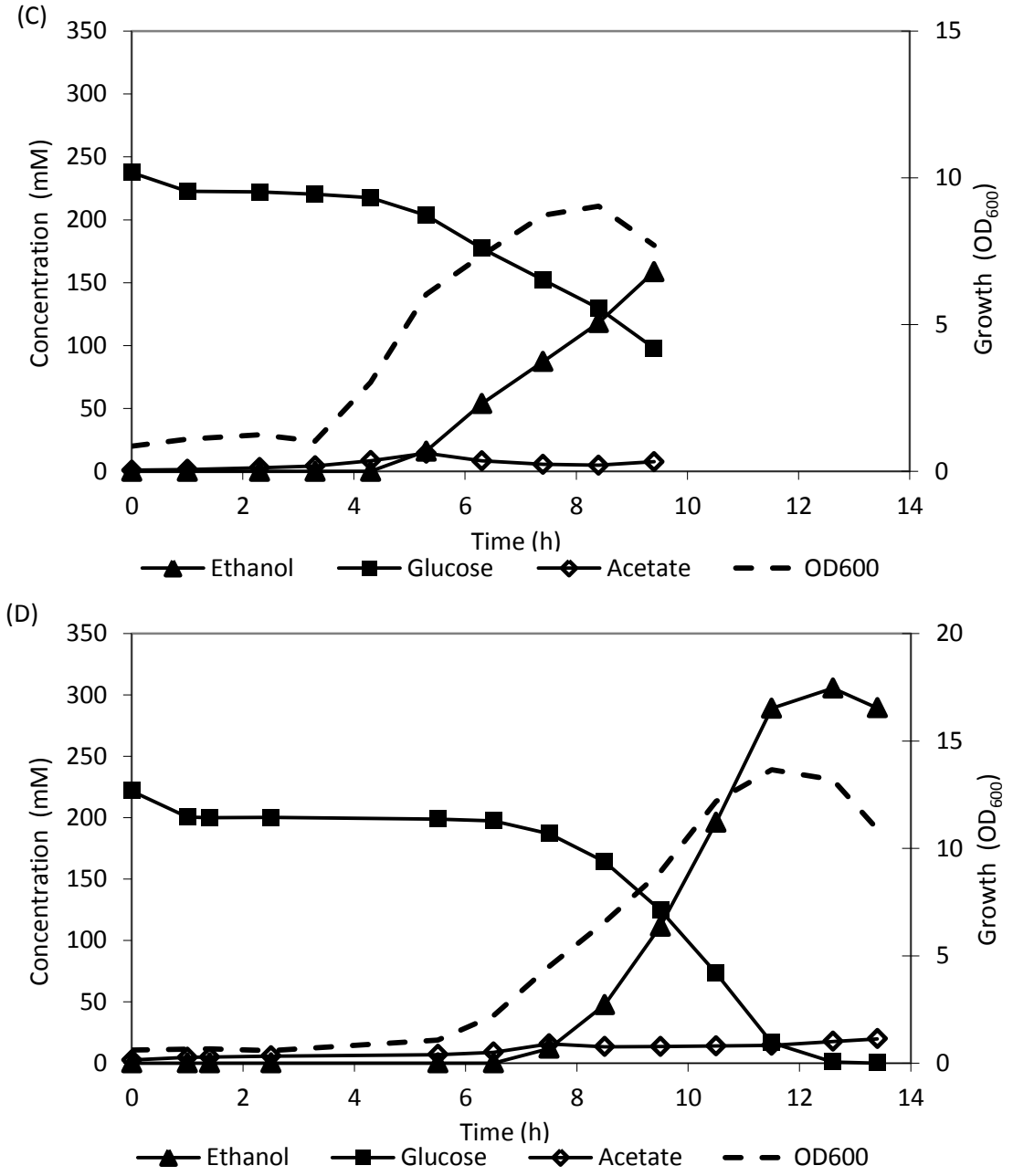


Figure 7.19 Products of fermentation by *G. thermoglucosidasius*. Cells were cultured on 4% (w/v) glucose in a USM medium containing 2% (w/v) yeast extract, as described in Section 2.5.2. The x-axis displays the time post inoculation, the primary y-axis displays the concentration of metabolites, and the secondary y-axis displays the growth in OD at 600nm. The anaerobic switch was performed at 4h for TM242 (A), 2h for Δ*pta* (B), 4.3h for Δ*ptp* (C) and 8h for Δ*ptb* (D).

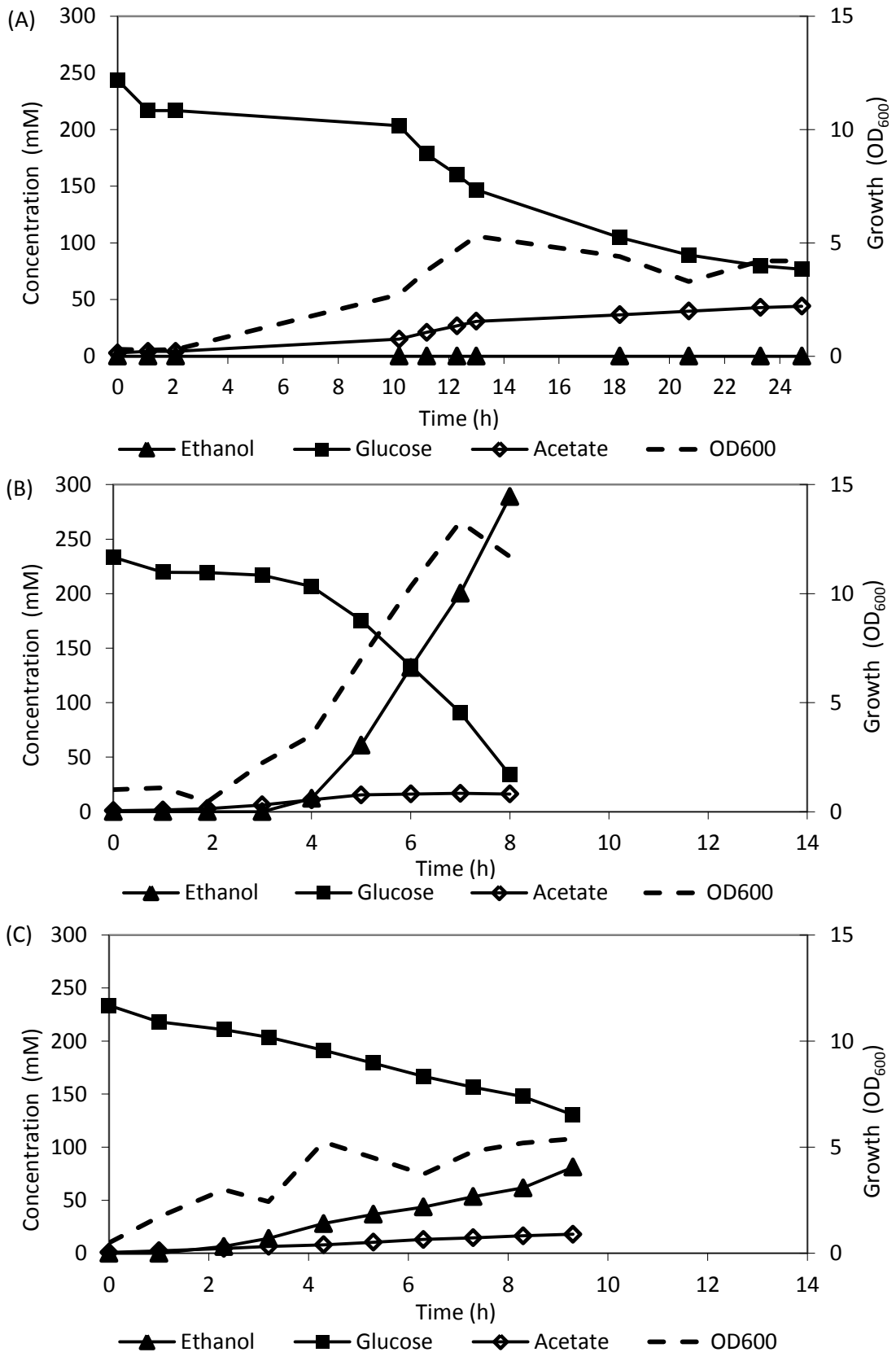


Figure 7.20 Products of fermentation by *G. thermoglucosidasius*. Cells were cultured on 4% (w/v) glucose in a USM medium containing 2% (w/v) yeast extract, as described in Section 2.5.2. The x-axis displays the time post inoculation, the primary y-axis displays the concentration of metabolites, and the secondary y-axis displays the growth in OD at 600nm. The anaerobic switch was performed at 13h for Δ*pta*Δ*ptb* (A), 5h for Δ*ptp*Δ*ptb* (B) and 4.3h for Δ*pta*Δ*ptp*Δ*ptb* (C).

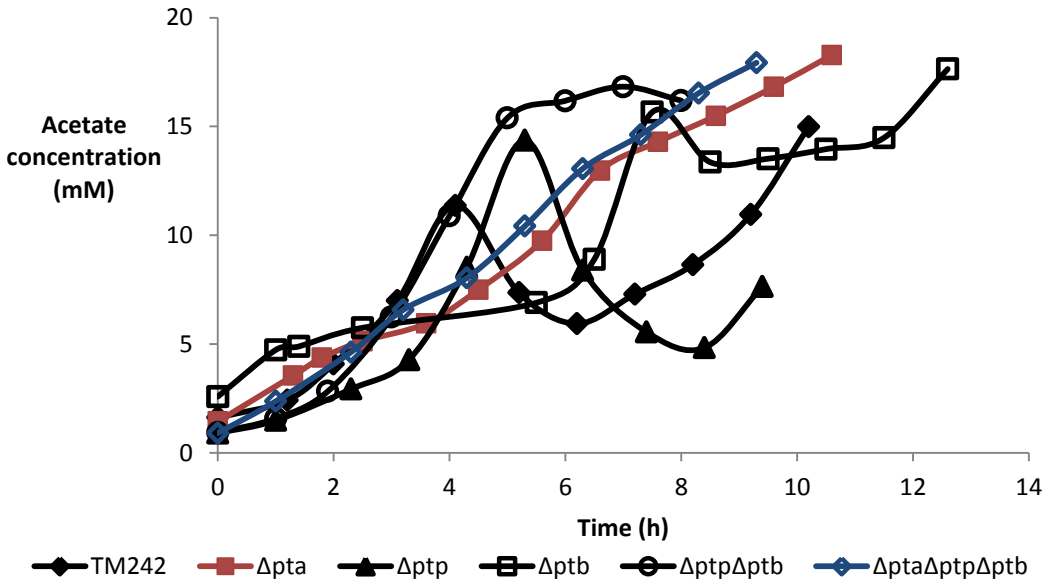


Figure 7.21 Acetate profiles for *G. thermoglucosidasius* batch-fermentations on USM medium containing 4% (w/v) glucose. The strains carrying deletions to *pta* are in red and blue to highlight the lack of acetate assimilation.

Batch fermentations were also run on USM containing 3% (w/v) xylose and 2% (w/v) yeast extract. However, the growth of many of the strains was very inconsistent and there were concerns over possible contamination to some of the cultures. Due to these factors the data are discussed here but are not fully analysed. The TM242 and Δ*pta* data have been analysed in Section 6.3.2.3, but it is worth noting that the TM242 strain did not seem typical because it had not consumed all the available xylose. Despite healthy seed growth, the Δ*ptp*, Δ*pta*Δ*ptp*, and Δ*ptp*Δ*ptb* strains did not grow in the fermenter tank. The Δ*ptb* strain twice failed to grow, but on one instance the strain grew but did not consume all the xylose and produced no ethanol and high levels of lactate and acetate. Due to the inconsistent data, questions over sterility were asked. Inconsistent data were also observed with the Δ*pta*Δ*ptb* strain; growth was observed in each tank but neither culture consumed all the available xylose. However, a significantly greater quantity of lactate was produced in one culture but not the second. The lack of a viable seed culture of the triple PTAC knockout prevented batch fermentations of these strain on xylose from being run.

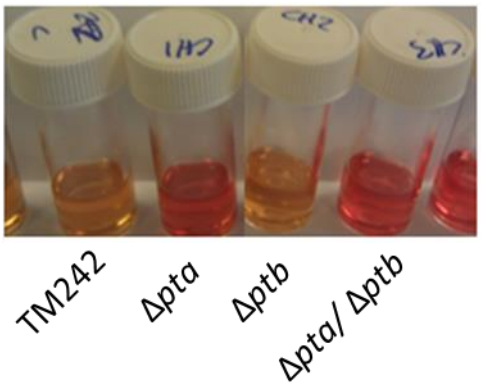


Figure 7.22 A qualitative test for acetoin production from *G. thermoglucosidasius* strains. Cultures were grown overnight in 2SPYNG and the presence of acetoin was tested using the Voges-Proskauer method. A yellow colour indicates a negative result, while a red colour indicates a positive result for the detection of acetoin.

7.4 Discussion

The work conducted in this Chapter demonstrated that the purified His-tagged PTP and PTB enzymes from *G. thermoglucosidasius* TM242 can catalyse the acyl-group transfer to phosphate from acetyl-CoA. Therefore the PTP and/or PTB could potentially compensate for the loss of PTA in the Δ*pta* strain, and could be involved in acetate formation. The specific activity determined for both PTP and PTB were similar (Table 7.3). This is around 5-fold lower than that determined for the truncated PTA enzyme but still well within the range of the specific activities of previously reported PTA enzymes (13 – 9100 U/mg BRENDA database). The concentration of orthophosphate required for both PTP and PTB to reach half $V_{\max}$ was almost identical. But the K_{m} for acetyl-CoA was 2.5-fold higher for PTP than that for PTB, suggesting that PTB is more efficient than PTP at catalysing the reaction between acetyl-CoA and phosphate.

Table 7.3 Kinetic parameters determined for truncated PTA, PTP and PTB. Truncated PTA parameters were determined in Chapter 4. Values are given, plus or minus standard errors.

Enzyme (substrate)	Substrate	$V_{\max}$ (U/mg)	K_{m} (mM)
Truncated PTA	acetyl-CoA	1,580 (±140)	0.059 (±0.017)
	phosphate	1,580 (±100)	3.5 (±0.9)
PTP	acetyl-CoA	269 (±11)	0.17 (±0.02)
	phosphate	250 (±19)	21 (±3) ($K_{0.5}$)
PTB	acetyl-CoA	260 (±66)	0.067 (±0.046)
	phosphate	319 (±19)	20 (±3)

To determine whether PTP and/or PTB were contributing to the catalysis of acyl-group transfer to phosphate from acetyl-CoA *in vivo*, the PTP and PTB encoding genes were deleted from *G. thermoglucosidasius* TM242. The Δ*ptp* strain had activity comparable to TM242 with acetyl-CoA, propionyl-CoA and butyryl-CoA, and the activities of the Δ*pta*Δ*ptp* strain with all three substrates were comparable to the Δ*pta* strain. The Δ*pta*Δ*ptb* strain had no activity with the three substrates. These data suggest that PTP is not expressed under the experimental conditions used, or that the enzyme is in an inactive state. These data also suggests that PTB is capable of catalysing the reaction with acetyl-CoA *in vivo*, and may be involved in acetate production in the Δ*pta* strain.

Despite the complete abolition of PTA-specific activity in the Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains, acetate was still generated. In fact, acetate yields increased under growth on a xylose containing medium, although the acetate yield of the triple PTAC knockout was 30% lower under growth on glucose compared to TM242. This would indicate that acetate is produced via an alternative pathway in the triple PTAC knockout strain.

Notable alterations to the acetate profile were observed with strains carrying deletions to the *pta* gene when these strains were cultured on a glucose-based medium (Figure 7.21). The Δ*pta* and the Δ*pta*Δ*ptp*Δ*ptb* strains did not re-assimilate any acetate generated. In contrast the TM242 acetate profile showed an initial increase in acetate, which subsequently dropped before acetate accumulation was observed again. A similar profile was also detected for the Δ*ptp* and Δ*ptb* strains. This 'acetate switch' has been extensively studied in *E. coli* (Alina Campos-Bermudez *et al.* 2010; Wolfe 2005; 2008) where, during exponential growth on a rich medium, acetate is excreted by *E. coli* as a way of regenerating CoA and NAD^+ pools. Then, during the transition to stationary phase, the cells re-assimilate the excreted acetate by either acetyl-CoA synthetase (ACS) or by reversal of the PTA-AK pathway. The PTA-AK pathway is only capable of utilising acetate at high concentrations. The inability to re-assimilate the acetate generated in *G. thermoglucosidasius* strains containing deletions to *pta* suggests the PTA-AK pathway is involved in the 'acetate switch' in *G. thermoglucosidasius*.

The Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains when fermenting xylose produced greater yields of pyruvate, lactate, citrate and acetate, compared to TM242. The secretion of pyruvate in particular indicates that these two strains lack the downstream ability to metabolise the

pyruvate generated. It is possible that the increased lactate yields are a result of the organism trying to balance its redox potential via a D-lactate dehydrogenase.

The Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains were capable of consuming all the available glucose, and despite there being some variability in the results the end-point products of fermentation were similar to those produced by the TM242 strain. However, these two deletion strains were poor growers on xylose, consuming less than 40% of the available sugar after 72h. In comparison, TM242 and all other PTAC knockouts consumed all the available xylose. The difference between glucose and xylose growth observed for the Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains could be attributed to the loss of ATP generated via the PTA-AK pathway. The ATP yield of glycolysis is 2 ATP molecules per molecule of glucose consumed, whereas the ATP yield of the pentose phosphate pathway is 1.67 molecules of ATP per molecule of xylose (see Figure 1.2 and Figure 1.3). The loss of the ATP generated via the PTA-AK pathway in the Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains could explain the inability to consume all the available xylose, and may suggest the ATP production ability of the TM242 strain is in a fine balance to support growth on both sugars. The 20-40% of the xylose consumed may have occurred under initial aerobic conditions, and growth and metabolism may arrest when the oxygen levels have dropped below a certain limit.

In summary, *G. thermoglucosidasius* TM242 has three genes that encode PTAC enzymes capable of using acetyl-CoA as a substrate. The promiscuous nature of PTP and PTB could potentially be used to compensate for the loss of PTA-specific activity in the Δ*pta* strain. A genetic knockout strategy has indicated PTB is more likely to be involved *in vivo*; however, the abolition of PTA-specific activity in the Δ*pta*Δ*ptb* and Δ*pta*Δ*ptp*Δ*ptb* strains did not prevent acetate production suggesting an alternative acetate-producing pathway exists. Despite this, it was discovered that a strain carrying a deletion to *pta* cannot assimilate the acetate produced under the conditions of the experiment in this Thesis. It would be interesting to observe whether these strains are capable of growth on a minimal medium containing acetate as the sole carbon source.

Chapter 8

Acetyl-CoA synthetase and regulation by lysine acetylation

8.1 Introduction

Prokaryotes often excrete acetate when sugar sources are in abundance; however, many are also able to utilise acetate as a carbon source. Extracellular acetate concentrations can vary widely, and prokaryotes have evolved three distinct pathways to convert acetate to acetyl-CoA to deal with the fluctuation in acetate concentration. At high concentrations (>30mM) the combination of acetate kinase (AK) and phosphotransacetylase (PTA) catalyses acetate to acetyl-CoA. At lower concentrations prokaryotes can use acetyl-CoA synthetase (ACS), of which there are two forms: an ADP-forming ACS (EC 6.2.1.1.3) and an AMP-forming ACS (EC 6.2.1.1). The ADP-forming ACS catalyses the reversible reaction as shown in Equation 8.1 in a single step. The AMP-forming ACS uses a different reaction mechanism, where it synthesises acetyl-CoA in two steps by first forming an acetyl-AMP intermediate as shown in Equation 8.2.

$$CH_3COOH + CoA + ATP \rightleftarrows CH_3COCoA + ADP + P_i \quad (8.1)$$

$$CH_3COOH + CoA + ATP \rightarrow \text{acetyl-AMP intermediate} \rightleftarrows CH_3COCoA + AMP + PP_i \quad (8.2)$$

The synthesis of acetyl-CoA by the AMP-forming ACS is catalysed via an ordered Bi Uni Uni Bi ping-pong mechanism (Starai and Escalante-Semerena 2004a). ATP is bound to the enzyme first, then acetate, which leads to the formation of enzyme-bound acetyl-AMP and the release of pyrophosphate. The CoA then binds to the enzyme, which catalyses the formation of acetyl-CoA before it releases acetyl-CoA followed by AMP. The formation of the acetyl-AMP intermediate is irreversible *in vivo* due to the formation of pyrophosphate, which would subsequently be cleaved by pyrophosphatases within the cell. Therefore, the AMP-forming ACS is considered to only be an acetate-uptake system *in vivo*.

A BLAST search of the *G. thermoglucosidasius* TM242 genome revealed four genes that are annotated as encoding putative ACS enzymes. The four putative ACS enzymes have a sequence identity of between 21% and 32% to the AMP-forming ACS enzymes from *Salmonella enterica* and *Saccharomyces cerevisiae*, and a sequence identity of between 19% and 35% to each other (Figure 8.1). This was significantly greater than the sequence identity to known ADP-forming ACS enzymes, suggesting that the ACS enzymes from

G. thermoglucosidasius belong to the AMP-forming family. The regions that are highly conserved amongst the AMP-family of ACS enzymes also suggest that the *G. thermoglucosidasius* ACS enzymes belong to the AMP-forming family (Figure 8.1) (Starai and Escalante-Semerena 2004a).

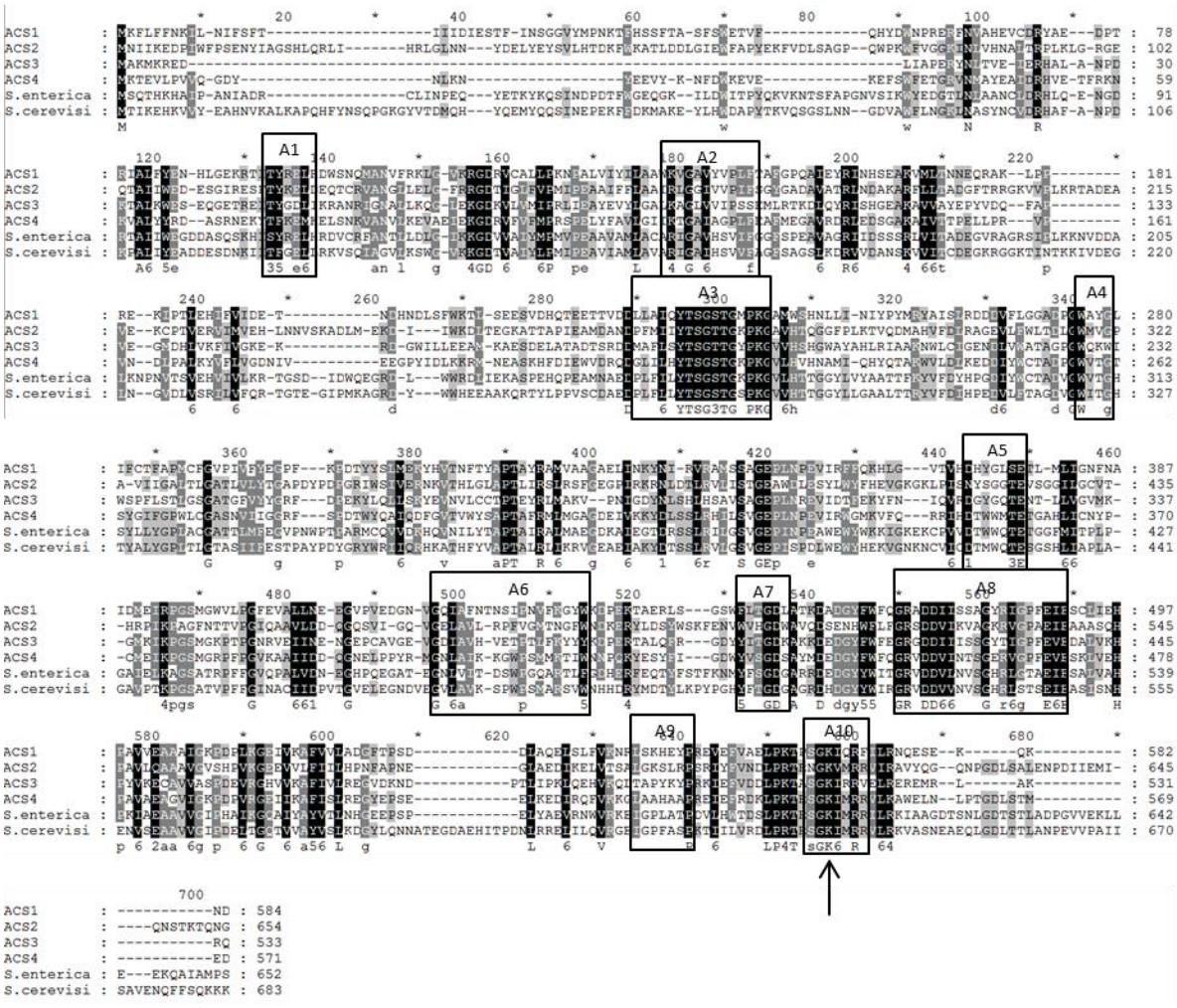


Figure 8.1 Identity amongst the *G. thermoglucosidasius* TM242 ACS enzymes and AMP-forming ACS enzymes of *Salmonella enterica* (UniProtKB accession number: V5S967) and *Saccharomyces cerevisiae* (UniProtKB accession number: P52910). The sequence motifs conserved amongst members of the AMP-forming ACS enzymes are boxed and labelled A1 to A10. The arrow indicates the lysyl residue of the A10 domain that is reversibly acetylated.

It is well documented that AMP-forming ACS enzymes are post-translationally regulated (Gardner *et al.* 2006; Starai *et al.* 2002; Starai and Escalante-Semerena 2004a, b). This regulation is via the acetylation/deacetylation of a lysyl residue of the A10 motif (Figure 8.1), which is essential for the activation of acetate to acetyl-AMP (Figure 1.4) (Starai *et al.* 2002). The acetylation of this lysine completely blocks acetyl-AMP synthesis, and is catalysed by an acetyl-CoA-consuming lysine acetyltransferase (abbreviated to LAT or KAT,

or sometimes known as protein acetyltransferase (PAT), but this is not to be confused with phosphotransacetylase (PTA)). There are three major families of KATs: (i) the MYST family; (ii) the Gcn5-related *N*-acetyltransferase (GNAT) family, and (iii) the CBP/p300 co-activators (Jones and O'Connor 2011). The re-activation of ACS is catalysed by lysine deacetylases (KDACs), which can be divided into two families: simple hydrolases which release the acetyl-group as acetate, and NAD^+-dependent deacetylases which cleave NAD^+ and generate nicotinamide and 2'-*O*-acetyl-ADP-ribose as co-products (Hu *et al.* 2010).

Specifically the ACS acetylation/deacetylation system in *S. enterica* involves a KAT and an NAD^+-dependent KDAC known as CobB (Starai *et al.* 2002; Starai and Escalante-Semerena 2004b). This suggests that there is a link between ACS activation (carbon metabolism) and NAD^+ biosynthesis (energy generation) in *S. enterica*. On the other hand, the ACS regulatory system in *Bacillus subtilis*, a Gram-positive bacterium closely related to *G. thermoglucosidasius*, involves a KAT that is a member of the GNAT family and a KDAC that is a hydrolase and therefore does not require NAD^+ (Gardner *et al.* 2006). The KAT and KDAC are expressed from a *B. subtilis* operon annotated as *acuABC* (acetoin utilisation), where the AcuA and AcuC were determined to be a KAT and KDAC, respectively. The AcuB is homologous to proteins with unknown function. BLAST analysis of the *G. thermoglucosidasius* TM242 genome identified an operon annotated as *acuABC*, encoding proteins with putative functions similar to those from *B. subtilis*.

Over the last decade it has become apparent that lysine acetylation occurs on a global scale in prokaryotes, whereas before this phenomenon was thought to be restricted only to eukaryotes (Hu *et al.* 2010; Jones and O'Connor 2011; Soppa 2010; Wang *et al.* 2010). Lysine acetylation plays an important role in the regulation of transcription, translation, the central metabolic pathways and the stress responses (Finkel *et al.* 2009; Zhao *et al.* 2010). A number of groups have identified lysyl residues on proteins that are acetylated by purifying acetylated peptides by the use of antibodies specific to acetyllysine and identifying those peptides by using mass spectrometry: 125 acetylated sites were identified in 85 different proteins in *E. coli* (Yu *et al.* 2008); 253 acetylated sites were identified in 191 different proteins in *S. enterica* (Wang *et al.* 2010); and 332 acetylated sites were identified in 185 different proteins in *B. subtilis* (Kim *et al.* 2013). In addition, 253 acetylated sites were discovered in 114 different proteins from *Geobacillus kaustophilus*, the only thermophile to have had its acetylproteome studied to date (Lee *et al.* 2013).

There is surmounting evidence to the importance of acetylation in prokaryotes, and it has been suggested that lysine acetylation is comparable to phosphorylation. However, despite this, the mechanism of acetylation largely remains unknown in prokaryotes. A small handful of enzymes are known to be directly regulated by the action of KATs and KDACs including ACS, a chemotaxis protein CheY and the response regulator RcsB (reviewed by Jones and O'Connor (2011)). However, considering hundreds of proteins are now known to be acetylated, there are only a comparatively small number of KATs and KDAC enzymes that have been identified.

In this Chapter the role of ACS in *G. thermoglucosidasius* TM242 is investigated, and the results suggest that acetylation could play a part in the regulation of its function. There is evidence that acetylation may also be involved in regulating the activity of phosphotransacetylase (PTA), which suggests acetylation may be a common post-translational modification in *G. thermoglucosidasius*. It was discovered that the acetylation of PTA is non-enzymatic but pH-dependent, suggesting that acetylation may be triggered as a response mechanism to changing environments. In this Chapter, kinetic characterisation of ACS4 and the acetylation studies were carried out by supervised undergraduate project students Mary Lunson and Nick Wright, respectively.

8.2 Methods

8.2.1 Cloning

Cloning of the *acs* and *acu* constructs used general molecular biology techniques described in Section 2.6. Vector-specific primers were used for DNA sequencing to determine the correct sequence had been cloned. pET-vector-based clones were transformed into *E. coli* BL21(DE3) cells for protein expression.

8.2.1.1 *acs* genes

The *acs* genes 2-4 were amplified by the polymerase chain reaction (PCR) from *G. thermoglucosidasius* TM242 genomic DNA, using the forward (F) and reverse (R) primers specific to each gene (Table 8.1). The PCR products were purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *acs2* gene was excised from the pGEM®-T vector by restriction digest with *Nde*I and *Xho*I, while the *acs3* and *acs4* genes

were excised from the pGEM®-T vector by restriction digest with *Nhe*I and *Xho*I. Gel-purified amplicons were ligated into pET-28a(+) vectors that had been linearized with the relevant restriction enzymes. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the constructs.

Attempts to clone *acs1* used both traditional ligation-based cloning and construction methods based on the Gibson isothermal enzymic DNA assembly method (Gibson *et al.* 2009). For the traditional cloning attempts, *acs1* was PCR amplified from *G. thermoglucosidasius* TM242 genomic DNA using the primer pairs *acs1F* and *acs1R*, and *acs1F_2* and *acs1R_2*. Each amplicon was digested with *Nde*I and *Xho*I, and ligation into both pGEM®-T and pET-28a(+) was attempted. For the Gibson assembly attempt, *acs1* was PCR amplified from *G. thermoglucosidasius* TM242 genomic DNA using primers *acs1GibF* and *acs1GibR*. The pET-28a(+) vector was digested with *Nde*I. The digested vector and amplicon were purified by gel electrophoresis, and the assembly of the construct was attempted using the Gibson isothermal enzymic DNA assembly method (Gibson *et al.* 2009).

8.2.1.2 *acu* genes

The *acuA* and *acuC* genes were PCR-amplified from *G. thermoglucosidasius* TM242 genomic DNA, using primer pairs *acuAF* and *acuAR*, and *acuCF* and *acuCR*, respectively (Table 8.1). The PCR products were purified, A-tailed, and ligated into the pGEM®-T easy vector. Ligated plasmids were transformed into *E. coli* JM109 cells, and blue/white screening was used to select positive colonies. The *acuA* gene was excised from the pGEM®-T vector by restriction digest with *Nhe*I and *Xho*I, agarose gel purified and ligated into pET-28a(+) that had been linearized with *Nhe*I and *Xho*I. The *acuC* gene was excised from the pGEM®-T vector by restriction digest with *Nde*I and *Bam*HI, agarose gel purified and ligated into pET-19b(+) that had been linearized with *Nde*I and *Bam*HI. Ligated plasmids were transformed into *E. coli* JM109 cells to amplify the constructs.

Table 8.1 PCR primers used for *acs* and *acu* cloning. Restriction sites are underlined, and regions homologous to pET-28a(+) for Gibson cloning are highlighted in grey.

Name	Sequence (5'-3')
acs1F	CATATGATGAAATTTTTGTTTTTTAATAAAATATT
acs1R	CCTCGAGGTTAGTCGTTTTTCTGTTTTTC
acs1F_2	CATATGATGAAATTTTTGTTTTTTAATAAAATATTGAATATT
acs1R_2	CCTCGAGGTTAGTCGTTTTTCTGTTTTTCGCTTTC
acs1GibF	GCCTGGTGCCGCGCGGCAGCCATATGATGAAATTTTTGTTTTTTAATAAAATATTGAATATT
acs1GibR	GATCCGCGACCCATTTGCTGTCCACCAGTACTGCTAGCTTAGTCGTTTTTCTGTTTTTCGCTTTC
acs2F	CATATGATGAACATCATAAAGGAAGATCC
acs2R	CCTCGAGGTTACCCATTTTGTGTTTTAGT
acs3F	GCTAGCATGGCGAAAATGAAACGGGAAGA
acs3R	CCTCGAGGTTATTGACGCTTGGCAAGGCG
acs4F	GCTAGCATGAAAACAGAAGTGCTACCGG
acs4R	CCTCGAGGTTAATCTTCCATTGTCGACAA
acuAF	GCTAGCATGAAAACAGAAGAAGAAATTCCAGG
acuAR	CTCGAGTTAATACATGTGCCGGTTCATG
acuCF	CATATGATGGCCGAATTTGCCGGGGGTTT
acuCR	GGATCCTTAGAAATGGTTGGACTGGTATCCGT

8.2.2 Protein purification

Recombinant proteins were purified by nickel-affinity chromatography as described in Section 2.7.4.

8.2.3 Enzyme characterisation

8.2.3.1 Acetyl-CoA synthetase activity

$$\text{acetate} + \text{CoA} + \text{ATP} \rightleftarrows \text{acetyl-CoA} + \text{AMP} + PP_i$$

Acetyl-CoA synthetase (ACS) activity was measured in the acetyl-CoA-forming direction by using one of three stopped assay systems. Each method involved a step where the ACS reaction took place, then the amount of product formed or substrate consumed was measured by using one of three assays. The initial ACS reaction mixture (500µl) typically contained 100mM MES, pH 6.5 at 60°C (unless otherwise stated), 10mM $MgCl_2$, 2mM ATP, 0.5mM CoA, 5mM acetate and enzyme. It is important to note that the stock CoA was

dissolved in 20mM citrate, pH 5.1 at 24°C to maintain stability. The reaction was allowed to proceed for a pre-determined period of time (typically 80s) before the reaction was stopped and the rate determined by one of three assays:

1. Malate dehydrogenase and citrate synthase coupled assay.

The initial ACS assay was stopped by the addition of 500µl of ice-cold stop buffer (200mM Tris-HCl, pH 7.6 at 25°C, 0.6mM NAD^+, and 5mM malate). The concentration of acetyl-CoA produced by the ACS was determined by measuring the absorbance at 340nm (ε_{340} = 6,200 $M^{-1}.cm^{-1}$) in a coupled assay with malate dehydrogenase and citrate synthase at room temperature. Typical reaction mixtures (500µl) contained 200mM Tris-HCl, pH 8.0 at 25°C, 20mM malate, pH 4.25 at 25°C, and 2.4U of malate dehydrogenase. Then 500µl of the stopped ACS reaction was added and the absorbance at 340nm was recorded. The change in absorbance at 340nm after the addition of 3U of citrate synthase was used to calculate the concentration of acetyl-CoA. The citrate synthase reaction goes to >99% completion, converting acetyl-CoA and oxaloacetate to citrate and CoA. This disrupts the equilibrium created by malate dehydrogenase, resulting in the catalysis of malate and NAD^+ to oxaloacetate and NADH + H^+. The increase in absorbance at 340nm due to NADH production was equivalent to the µmoles of acetyl-CoA formed by ACS.

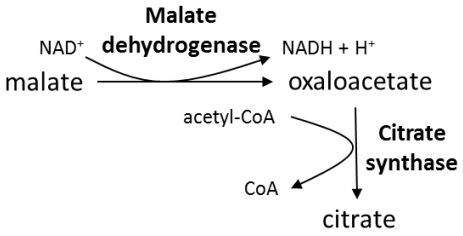


Figure 8.2 Schematic representation of the malate dehydrogenase and citrate synthase coupled assay system used to measure ACS activity. The acetyl-CoA production by acetyl-CoA synthetase was coupled to the activities of malate dehydrogenase and citrate synthase and the rate of NADH production was measured at 340nm.

2. Loss of CoA with DTNB.

The initial ACS assay was stopped by the addition of 500µl of ice-cold stop buffer containing DTNB (200mM Tris-HCl, pH 7.6 at 25°C, and 2mM DTNB). The concentration of CoA in the stopped mixture was calculated by measuring the absorbance change at 412nm (ε_{412} = 13,600 $M^{-1}.cm^{-1}$) when 200µl of the stopped mixture was added to 800µl of 200mM Tris-HCl, pH 8.0 at 25°C. The loss of CoA was determined by calculating the difference in CoA concentration against an initial ACS reaction containing no enzyme. The rate of ACS was

determined as the µmoles of acetyl-CoA formed (equivalent to the µmoles of CoA consumed) per minute per mg of protein.

3. Citrate synthase coupled assay.

The initial ACS assay was stopped by the addition of 500µl of ice-cold stop buffer containing DTNB (200mM Tris-HCl, pH 7.6 at 25°C, and 2mM DTNB), or a sample of the stopped assay from the loss of CoA with DTNB method was used. The concentration of acetyl-CoA in the stopped mixture was calculated by measuring the absorbance change at 412nm (ε_{412} = 13,600 $M^{-1}.cm^{-1}$) in a coupled assay with citrate synthase at room temperature. Typical reaction mixtures (1ml) contained 200mM Tris-HCl, pH 8.0 at 25°C, 100µl of the stopped mixture, and 3U of citrate synthase. The change in absorbance at 412nm after the addition of 20mM oxaloacetate was added was used to calculate the concentration of acetyl-CoA. The rate of ACS was determined as the µmoles of acetyl-CoA formed per minute per mg of protein.

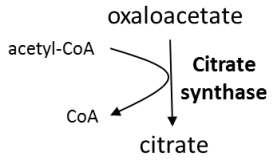


Figure 8.3 Schematic representation of the citrate synthase coupled assay system used to measure ACS activity. The acetyl-CoA production by acetyl-CoA synthetase was determined by measuring the increase in A_{412nm}, corresponding to the released CoA reacting with DTNB.

8.2.4 Acetyl-CoA synthetase regulation

The effects of AcuA and AcuC on ACS4 were determined using purified recombinant proteins. The purified ACS4 was dialysed overnight into a 50mM Hepes buffer, pH 6.5 at 25°C, while AcuA and AcuC were dialysed into 50mM Tris-HCl, pH 8.0 at 25°C. The regulatory role of AcuA was determined by incubating 30µl of ACS4 with 20µl of AcuA and 10µl of 0.07mM acetyl-CoA (in 20mM citrate, pH 5.0), for 10min at 60°C. When ACS4 was incubated in the absence of AcuA or acetyl-CoA, 20µl of 50mM Tris-HCl, pH 8.0 or 10µl of 20mM citrate, pH 5.0, respectively, was added to maintain the buffer composition. The activity of ASC4 was determined by the loss of CoA method as described in Section 8.2.3.1.

The regulatory role of AcuC was determined by incubating 30µl of ACS4 with 20µl of AcuC and 10µl of 0.07mM NAD^+ (in 50mM MES, 10mM $MgCl_2$ pH 6.5), for 10min at 60°C. When

ACS4 was incubated in the absence of AcuC or NAD^+, 20µl of 50mM Tris-HCl, pH 8.0, or 10µl of 50mM MES, 10mM $MgCl_2$ pH 6.5, respectively, was added to maintain the buffer composition. The activity of ASC4 was determined by the loss of CoA method as described in Section 8.2.3.1.

8.2.5 The acetylation of phosphotransacetylase

The acetylation of PTA was investigated using truncated PTA, purified from *E. coli* BL21(DE3) cells, and from native *G. thermoglucosidasius* TM242 cell lysate. The PTA sample was dialysed overnight into 50mM Tris-HCl, at either pH 6.5 or 8.0 at 25°C. AcuA was dialysed into 50mM Tris-HCl, pH 8.0 at 25°C. The PTA was diluted into a 50mM Tris-HCl buffer, at the corresponding pH, in the presence or absence of AcuA and/or 0.07mM acetyl-CoA. This mixture was incubated at 60°C for 10min before the PTA activity was measured as described in Section 2.7.7.3.

8.3 Results

8.3.1 Cloning and recombinant expression

Of the four *acs* genes identified, three were successfully cloned into the *E. coli* expression vector pET-28a(+). Several attempts were made to clone *acs1*, including by traditional ligation cloning and cloning based on the Gibson isothermal assembly method (Gibson *et al.* 2009). However, a successful clone of *acs1* was not created by either method. The *acs* genes 2-4 were expressed in *E. coli* BL21(DE3) cells using the inducible T7 expression system. This resulted in recombinant N-terminal His-tagged proteins with molecular masses close to the predicted molecular masses of 72.4kDa, 60.2kDa and 65.2kDa for ACS2, ACS3 and ACS4 respectively. Proteins were purified by nickel affinity chromatography and visualised by SDS-PAGE (Figure 8.4).

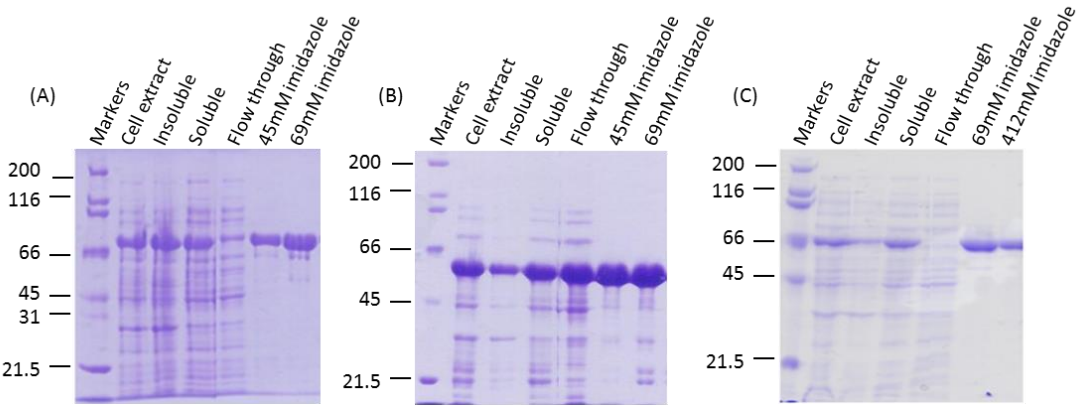


Figure 8.4 SDS-PAGE analysis of recombinant His-tagged ACS proteins. His-tagged ACS2 (A), ACS3 (B), and ACS4 (C) with expected molecular masses of 72.4kDa, 60.2kDa and 65.2kDa, respectively, were purified from *E. coli* BL21(DE3) cell lysates by nickel-affinity chromatography. The soluble fraction of each cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and the proteins were purified by washes with increasing imidazole concentrations. The sizes of the marker proteins are in kDa.

The genes encoding the ACS regulatory proteins AcuA and AcuC were cloned by PCR-amplification from *G. thermoglucosidasius* TM242 genomic DNA. The *acuA* and *acuC* genes were sub-cloned into pET-28a(+) and pET-19b(+) respectively. The AcuA and AcuC proteins were expressed in *E. coli* BL21(DE3) cells using the inducible T7 expression system. This resulted in recombinant His-tagged proteins with molecular masses close to the predicted molecular mass of 28.3kDa and 48.1Da for AcuA and AcuC, respectively. Proteins were purified by nickel-affinity chromatography and visualised by SDS-PAGE (Figure 8.5).

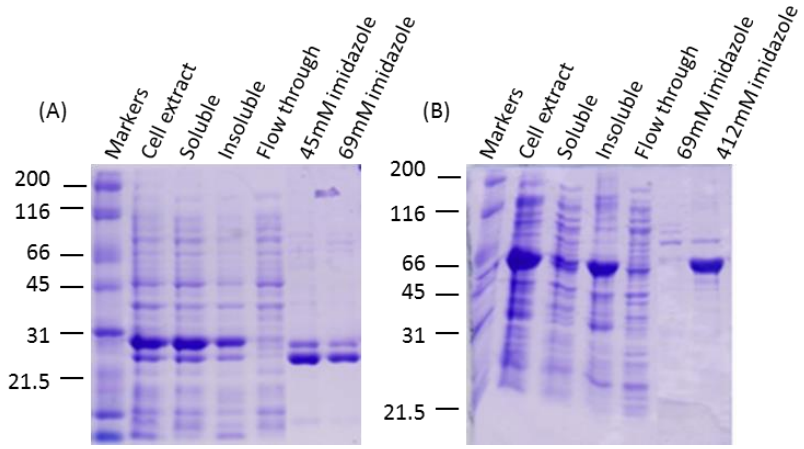


Figure 8.5 SDS-PAGE analysis of recombinant His-tagged ACS regulatory proteins. His-tagged AcuA (A), and AcuC (B) with expected molecular masses of 28.3kDa and 48.1kDa, respectively, were purified from *E. coli* BL21(DE3) cell lysates by nickel-affinity chromatography. The soluble fraction of each cell extract was separated from the insoluble debris. Soluble samples were loaded onto the column and each protein was purified by washes with increasing imidazole concentrations. The sizes of the marker proteins are in kDa.

8.3.2 Characterisation of acetyl-CoA synthetases

The purified recombinant ACS enzymes were assayed in the acetyl-CoA forming direction. No activity was observed with ACS2, and therefore no further studies were conducted on this protein. Activity was observed with both ACS3 and ACS4, and in both cases a linear relationship between the rate of catalysis and the amount of protein in the assay was evident. This suggests that the rate limiting factor of the assay was the amount of enzyme present in the assay and that initial rates were being measured.

Further characterisation of the ACS enzymes included determining the kinetic properties. The kinetic parameters for ACS3 were determined by coupling the stopped assay to citrate synthase activity. The rate of ACS3 displayed typical Michaelis-Menten kinetics towards acetate concentration (Figure 8.6 A and B) and yielded a $V_{\max}$ of 2.0 (± 0.1) U/mg and a $K_{\mathrm{m(acetate)}}$ of 6.0 (± 0.7) mM. What appeared to be substrate inhibition was observed with CoA, and so the data were fitted to the substrate inhibition rate equation (Figure 8.6 C and D), yielding a $V_{\max}$ of 3.1 (± 0.1) U/mg, a $K_{\mathrm{m(CoA)}}$ of 0.38 (± 0.02) mM, and a K_{i} of 1.4 (± 0.1) mM. Substrate inhibition was also observed with ATP, and these data were also fitted to the substrate inhibition rate equation (Figure 8.6 E and F), yielding a $V_{\max}$ of 2.0 (± 0.2) U/mg, a $K_{\mathrm{m(ATP)}}$ of 0.29 (± 0.05) mM, and a K_{i} of 10 (± 5) mM. The K_{i} for ATP is 33-fold higher than the K_{m} for ATP, suggesting that only weak substrate inhibition is observed with this substrate.

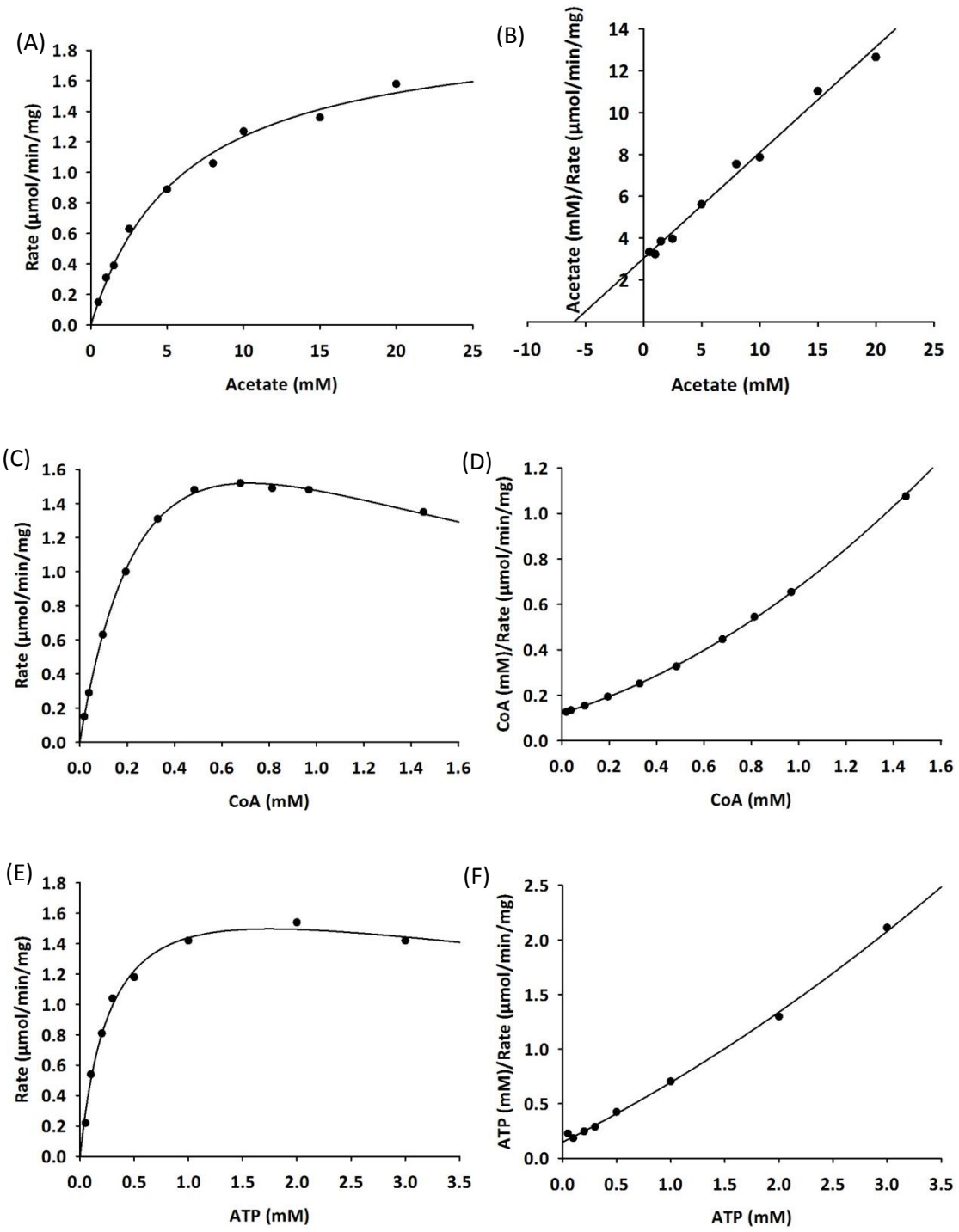


Figure 8.6 Acetyl-CoA synthetase 3 activity in the acetyl-CoA forming direction. The relationship between specific activity (U/mg) and acetate concentration (A and B); CoA concentration (C and D); ATP concentration (E and F) are displayed as Michaelis-Menten and Hanes-Woolf plots.

The kinetic parameters for ACS4 were determined by coupling the stopped assay to citrate synthase activity. What appeared to be substrate inhibition was observed with acetate, therefore the data were fitted to the substrate inhibition rate equation (Figure 8.7 A and B) and yielded a $V_{\max}$ of 10 (± 1) U/mg, a $K_{\mathrm{m(acetate)}}$ of 0.11 (± 0.02) mM, and a K_{i} of 36 (± 16) mM. What appeared to be substrate inhibition was observed with CoA, however, the data

could not be fitted to the inhibition rate equation so kinetic parameters could not be generated (Figure 8.7 C and D). Substrate inhibition was also observed with ATP; the data were fitted to the substrate inhibition rate equation (Figure 8.7 E and F), yielding a $V_{\max}$ of 22 (± 3) U/mg, a $K_{\mathrm{m(ATP)}}$ of 0.49 (± 0.13) mM, and a K_{i} of 2.9 (± 1.0) mM.

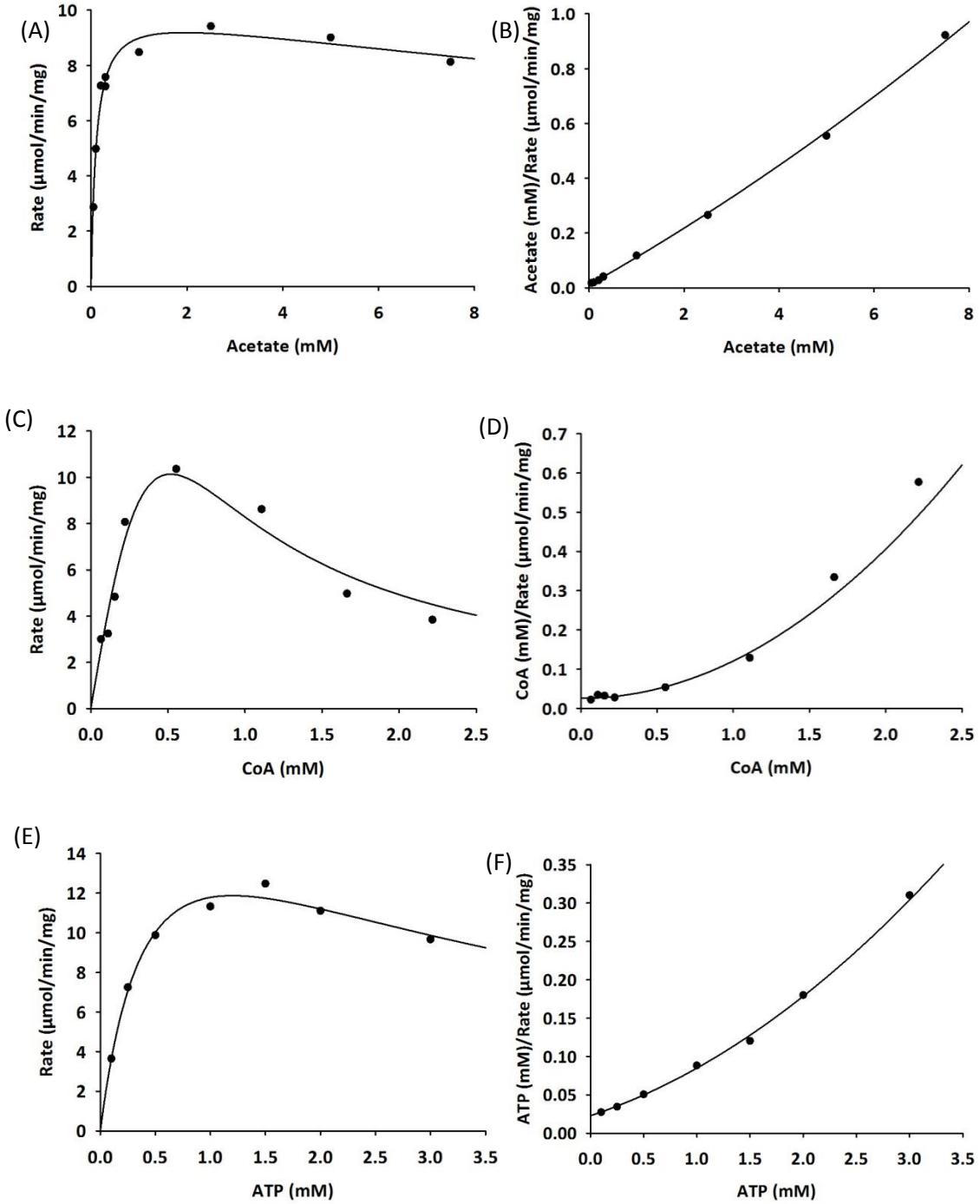


Figure 8.7 Acetyl-CoA synthetase 4 activity in the acetyl-CoA forming direction. The relationship between specific activity (U/mg) and acetate concentration (A and B); CoA concentration (C and D); ATP concentration (E and F) are displayed as Michaelis-Menten and Hanes-Woolf plots.

8.3.3 Regulation of acetyl-CoA synthetase

No ACS activity could be detected in *G. thermoglucosidasius* cell lysates, suggesting that an ACS was not expressed under the growth conditions of the experiment, or that any ACS present was in an inactive form. The *Bacillus subtilis* ACS is regulated by acetylation/deacetylation, which is controlled by the activities of AcuA and AcuC, respectively. The ACS is inactivated by acetylation with acetyl-CoA via AcuA, while AcuC reactivates ACS by deacetylation in an NAD^+-independent manner. Genes encoding proteins homologous to AcuA and AcuC from *B. subtilis* were found in the *G. thermoglucosidasius* TM242 genome. Here the regulatory role of *G. thermoglucosidasius* TM242 AcuA and AcuC was investigated by studying the effects of the purified recombinant proteins on purified ACS.

The acetylated state of the recombinant ACS enzymes purified from *E. coli* cell lysate is not known. However, activity was detected with ACS3 and ACS4 suggesting some if not all of the protein in the sample was in an active state (non-acetylated). The potential regulatory role of AcuA was investigated with ACS4 (Figure 8.8). ACS4 was incubated in 50mM Hepes, pH 6.5 at 60°C, for 10min with or without the presence of AcuA and/or acetyl-CoA. Then the ACS activity was determined using the loss of CoA assay method described in Section 8.2.3.1. No significant decrease in activity was observed when ACS4 was incubated with AcuA and acetyl-CoA. A decrease of 40% and 100% was observed when ACS4 was incubated with either acetyl-CoA or AcuA, respectively (Figure 8.8).

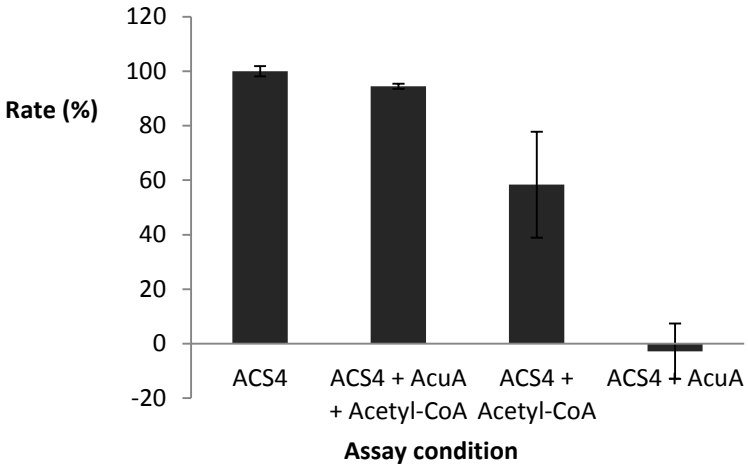


Figure 8.8 Regulation of acetyl-CoA synthetase 4 from *G. thermoglucosidasius* by AcuA. ACS4 was incubated for 10min in the presence and absence of AcuA and acetyl-CoA prior to detecting the activity of ACS4 in the acetyl-CoA forming direction.

To investigate whether AcuC could activate ACS4, the purified ACS4 protein was incubated under the conditions necessary to promote deacetylation (Figure 8.9). ACS4 was incubated in 50mM Hepes, pH 6.5 at 60°C, for 10min with or without the presence of AcuC and/or NAD^+. Then the ACS activity was determined using the loss of CoA assay method described in Section 8.2.3.1. An increase in activity was measured when ACS4 was incubated for 10min with AcuC alone, and AcuC with NAD^+, suggesting AcuC is an NAD^+-independent deacetylase. The incubation of ACS with NAD^+ alone resulted in a decrease of 27%.

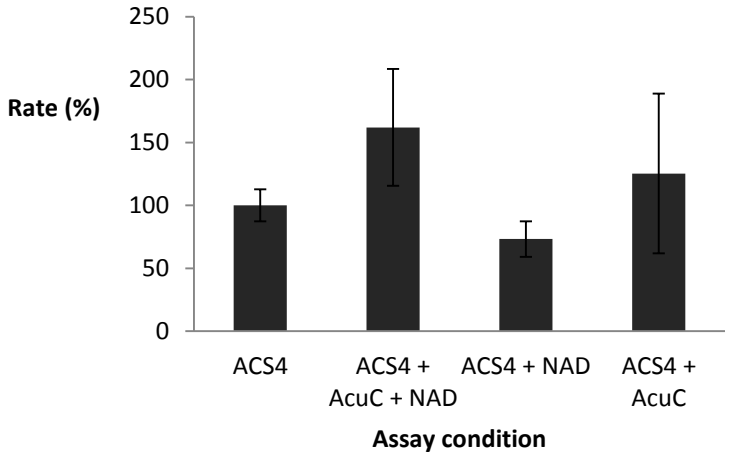


Figure 8.9 Regulation of acetyl-CoA synthetase 4 from *G. thermoglucosidasius* by AcuC. ACS4 was incubated for 10min in the presence and absence of AcuC and/or NAD^+ prior to detecting the activity of ACS4 in the acetyl-CoA forming direction.

8.3.4 The regulation of phosphotransacetylase by lysine acetylation

Lysine acetylation in prokaryotes is a relatively recent discovery, although it is now thought that prokaryotic enzymes are regulated by acetylation on a global scale (Wang *et al.* 2010). A study used mass spectrometry to identify that Lys^{250} from *B. subtilis* PTA was acetylated, and that this enzyme was one of 185 proteins to be acetylated in this organism (Kim *et al.* 2013). The sequence alignment against the *B. subtilis* PTA revealed that lysyl residues were conserved at positions 250 and 251 in the *G. thermoglucosidasius* TM242 PTA (Figure 8.10). It was hypothesised that lysine acetylation may have a role in regulating the *G. thermoglucosidasius* TM242 PTA. Here this hypothesis and the potential KAT and KDAC activity of AcuA and AcuC on PTA were investigated.

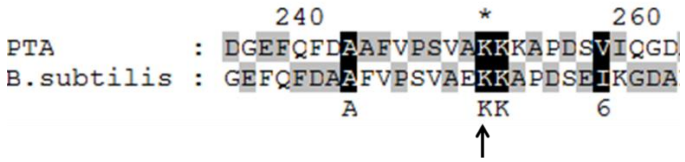


Figure 8.10 Homology between the PTAs from *G. thermoglucosidasius* TM242 (PTA) and *Bacillus subtilis*. The Lys^{251} residue that is reported to be acetylated in the *B. subtilis* PTA is indicated by an arrow.

The potential regulation of *G. thermoglucosidasius* TM242 PTA was investigated by incubating the purified recombinant truncated PTA with AcuA and acetyl-CoA (Figure 8.11). The PTA incubated in 50mM Tris-HCl at pH 8.0 in the presence of acetyl-CoA and AcuA resulted in a decrease in activity by over 60%; a similar decrease is observed when PTA is incubated with acetyl-CoA alone. In contrast the activity remained relatively constant when the incubation was repeated at pH 6.5. These results suggest that the acetylation of *G. thermoglucosidasius* PTA occurs by a non-enzymatic mechanism and is pH dependent. The small increase when PTA was incubated with acetyl-CoA alone at pH 6.5 could be attributed to the increase of acetyl-CoA concentration in the subsequent PTA activity assay, or could be an effect of 'priming' the PTA active sites with the acetyl-CoA substrate prior to measuring initial catalytic rates. This would also suggest that the PTA activity from samples incubated at pH 8.0 with acetyl-CoA is artificially high.

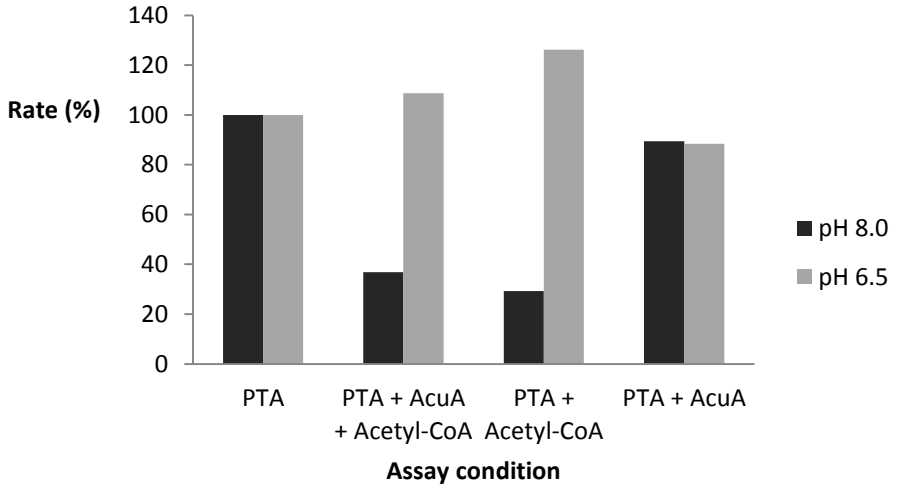


Figure 8.11 Regulation of PTA from *G. thermoglucosidasius* by AcuA. Truncated PTA was incubated for 10min in the presence and absence of AcuA and/or acetyl-CoA either at pH 8.0 or 6.5. Then the activity of PTA was measured in the acetyl-phosphate forming direction. The rate is given as a percentage of the rate of the PTA incubated alone.

A similar result was observed with the PTA activity from a *G. thermoglucosidasius* TM242 cell lysate. The PTA activity decreased by 40% when the cell lysate was incubated with

0.07mM acetyl-CoA at pH 8.0 but not at pH 6.5 (Figure 8.12). This again suggests that PTA could be regulated in *G. thermoglucosidasius* by non-enzymatic acetylation. Furthermore, PTA activity was reduced by 20% when as little as 0.018mM of acetyl-CoA was incubated with the cell lysate at pH 8.0, and the maximum reduction of 40% was observed with 0.035mM acetyl-CoA (Figure 8.13). A linear increase in PTA activity was observed when increasing concentrations of acetyl-CoA were incubated with cell lysate at pH 6.5. This increase in activity could be due to the increase of acetyl-CoA in the subsequent activity assay, or a 'priming' effect of PTA as described above.

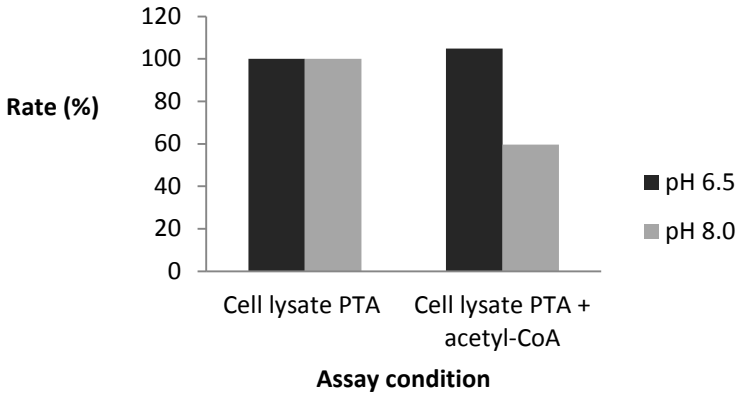


Figure 8.12 Regulation of phosphotransacetylase from *G. thermoglucosidasius* cell lysate by acetyl-CoA. Aliquots of *G. thermoglucosidasius* cell lysate were incubated for 10min in the presence and absence of acetyl-CoA either at pH 8.0 or 6.5. Then the activity of PTA was measured in the acetyl-phosphate forming direction. The rate is given as a percentage of the rate of the cell lysate PTA incubated in the absence of acetyl-CoA.

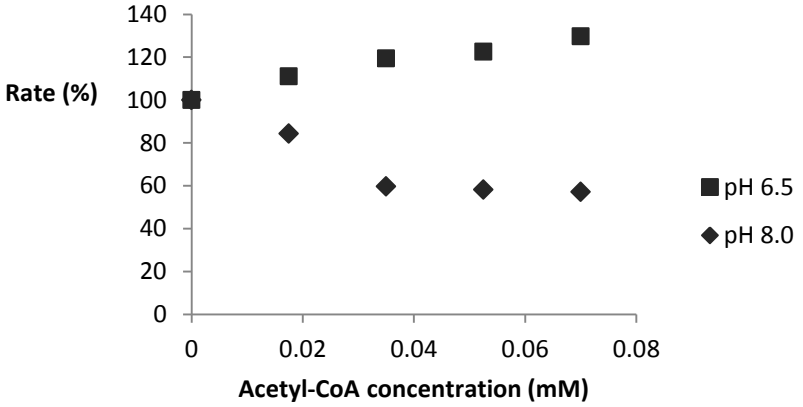


Figure 8.13 Acetyl-CoA regulation of phosphotransacetylase activity from *G. thermoglucosidasius* TM242 cell lysate. Aliquots of *G. thermoglucosidasius* cell lysate were incubated for 10min with different concentrations of acetyl-CoA, either at pH 8.0 or 6.5. Then the activity of PTA was measured in the acetyl-phosphate forming direction. The rate is given as a percentage of the rate of the cell lysate PTA incubated in the absence of acetyl-CoA.

8.4 Discussion

In this Chapter the *G. thermoglucosidasius* TM242 acetyl-CoA synthetase (ACS) enzymes were characterised. Recombinant expression of two of the four genes annotated in the genome generated functional enzymes, but no synthetase activity could be detected with ASC2; the gene encoding ACS1 was not cloned and therefore not tested. Although repeat data sets would improve the reliability, it was evident that ASC4 had approximately 10-fold higher $V_{\max}$ than ACS3, and a lower K_{m} for acetate. A similar observation has been made in *Saccharomyces cerevisiae*, which has two genes that encode ACS enzymes, one of which has a 30-fold lower K_{m} for acetate (van den Berg *et al.* 1996). Interestingly, the *S. cerevisiae* ACS enzymes are differentially expressed and one could use propionate as a substrate. The substrate specificity of the ACS enzymes has not been studied here, but one could propose *G. thermoglucosidasius* TM242 has four different ACS enzymes that could have varying degrees of activity with propionate, butyrate or other short chain organic acids. This may explain why no activity was measured with ACS2 with acetate.

An operon annotated as *acuABC* in *G. thermoglucosidasius* TM242 suggested that the ACS enzymes may be regulated by AcuA, a lysine acetyltransferase (KAT), and AcuC, a lysine deacetylase (KDAC), similar to the acetylation/deacetylation system of *B. subtilis* (Gardner *et al.* 2006). The presence of AcuC increased the activity of ACS4, and this increase was similar whether NAD^+ was present or absent, suggesting that NAD^+ was not required. The activity of ACS4 decreased in the presence of AcuA, however the addition of acetyl-CoA prevented this decrease. This is surprising as it is expected that acetyl-CoA is used by AcuA to acetylate and inhibit ACS. This result suggests that AcuA alone, or with another molecule present in the assay, can decrease the activity of ACS4. The use of a stopped assay to measure changes in ACS activity increases the chance of errors, and therefore repeating the experiment would increase the reliability of the data. In addition, the extent of ACS acetylation could be quantified using anti-acetyllysine immunoblot assays. A further note is that the potential regulation of ACS1, ACS2 or ACS3 by AcuA/AcuC has not been investigated here.

It is now known that lysine acetylation is a global event and occurs in many prokaryotes. However, there is little evidence to the mechanism of this acetylation in prokaryotes. The truncated PTA was inactivated when incubated with acetyl-CoA in alkaline conditions, and did not require AcuA or any other enzyme. This non-enzymatic acetylation of PTA is likely

being stimulated under these conditions, because at an alkaline pH the lysyl residues will be partially deprotonated and the lone pair of electrons on the N atom enable it to act as a nucleophile, attacking the C=O bond of acetyl-CoA. Further investigations into the conditions that stimulate the acetylation of PTA are still required to assess whether this could be used as a regulatory system under physiologically relevant conditions. However, this does provide a possible regulatory mechanism to control the carbon flux from acetyl-CoA by PTA in *G. thermoglucosidasius*. Furthermore, additional experiments are required to investigate whether non-enzymatic acetylation is a regulatory mechanism used by *G. thermoglucosidasius* on a global scale.

Kim *et al*. (2013) identified peptide fragments from *B. subtilis* that contained an acetylated Lys residue, and suggested Lys^{250} from PTA as a target for acetylation. A separate study crystallised the *B. subtilis* PTA, and suggested Lys^{251} is a critical catalytic residue because it was in a position from which it could form hydrogen bonds between the side chain amide group and the carbonyl oxygen of the acetyl-phosphate, and thus Lys^{251} would be involved in orientating the substrate (Xu *et al.* 2005). Considering that the mass spectroscopy approach used by Kim *et al.* (2013) would not be able to determine which lysyl residue was acetylated, if more than one Lys were on the peptide fragment, and that acetylation disrupted the activity, it is likely that Lys^{251} is the residue that is acetylated. This could be the case for the *G. thermoglucosidasius* PTA, but additional experiments would be necessary to determine this. A mutagenesis study could be used to identify which Lys residue is acetylated in the *G. thermoglucosidasius* PTA.

The results of this Chapter begin to suggest that acetylation could have a potentially significant role to play in regulating enzyme activity in *G. thermoglucosidasius*. However, many questions still remain: (i) are there other KATs and/or KDACs expressed by *G. thermoglucosidasius*? (ii) how many enzymes are regulated by acetylation? (iii) is the majority of acetylation enzymatically catalysed or non-enzymatic? (iv) is the acetylation of PTA reversible? (v) are there enzymes that can catalyse the acetylation/deacetylation of PTA? With so many questions arising from this study, and of similar studies, there is no doubt that the continued investigation of prokaryotic acetylation will bring some interesting observations that may well influence how we view metabolism as a whole in the future.

Chapter 9
General conclusion

Geobacillus thermoglucosidasius 11955 has been identified as a potentially useful candidate for bioethanol production from lignocellulosic sources. This was based on a multitude of factors including the organism's ability to degrade a wide range of substrates, such as hexose and pentose sugars, and some polycarbohydrates, in addition to possessing the genes required to produce ethanol. The engineered strain TM242 (Δ*ldh*, Δ*pfl*, *pdh*^{up}) has been demonstrated to produce ethanol in yields in excess of 90% of the theoretical maximum (Cripps *et al.* 2009). The second major product of fermentation is acetate, which is undesirable in a commercial process because it not only diverts the carbon flux away from ethanol, but it also acidifies the medium leading to an arrest in cellular growth. This Thesis, therefore, aimed to study and understand the metabolic pathways surrounding acetate production in *G. thermoglucosidasius* TM242, and to investigate potential methods of reducing or eliminating its production.

To model the metabolic flux in *G. thermoglucosidasius* TM242, an understanding of the complex relationship between the cellular regulation of enzyme properties, the environment, and the *in vivo* metabolite concentrations is required (Gerosa and Sauer 2011). Multiple factors are involved, including those that influence enzyme concentration and kinetic properties, such as gene expression and post-translational modification. Changes to the intracellular substrate concentration in response to environmental cues will also alter the reaction rate.

The activities of the bifunctional enzyme aldehyde dehydrogenase/alcohol dehydrogenase (ADHE), phosphotransacetylase (PTA), acetate kinase (AK) and the pyruvate dehydrogenase complex (PDH) from *G. thermoglucosidasius* TM242 cell pellets collected throughout a glucose fermentation were analysed in Chapter 3. These data were used not only to compare the relative activities of each enzyme but also to monitor changes in activity levels throughout the fermentation. The results indicated that the maximum ethanol production rate coincided with the maximum activity level of ADHE, suggesting that the *in vivo* concentration of ADHE and/or the regulation of its activity may be significant factors controlling the carbon flux to ethanol. The activity of PTA was more constant, however, and AK levels decreased throughout the fermentation. The acetate concentration steadily

increased during the fermentation, and there was no obvious maximum rate of acetate production that correlated to maximum concentrations of PTA and/or AK. This suggests that another factor may control the flux through to acetate. Considering that there is a significant increase in acetate production when *G. thermoglucosidasius* TM242 ferments xylose, and the discovery of sequences analogous to the *cre* sites involved in carbon catabolite regulation in *Bacillus subtilis*, it would be interesting to compare the expression levels of these enzymes from fermentations on these different sugars. This work, however, has not been conducted in this Thesis.

The absolute activities of PTA and AK were significantly greater than that for ADHE throughout the fermentation and, although consideration of assay conditions is necessary, these results suggest the potential flux is biased towards acetate production as opposed to ethanol. The kinetic parameters determined for PTA and AK from both purified recombinant proteins and native cell lysates reveal that these two enzymes have a significantly higher $V_{\max}$ than the activities of ADHE. The higher $V_{\max}$ and the fermentation activity level data suggest that the flux from acetyl-CoA has a greater potential towards acetate production as opposed to ethanol. However, the major fermentation product is ethanol, which indicates that other factors such as the regulation of enzyme properties, environmental conditions and substrate-product concentrations must be controlling the *in vivo* flux.

The aldehyde dehydrogenase (AldDH) domain of the ADHE and the PTA have similar K_{m} values for acetyl-CoA at 0.052mM and 0.059mM, respectively. A similar Michaelis constant indicates that a similar concentration of acetyl-CoA is required for each enzyme to reach half $V_{\max}$, and suggests that the affinity for acetyl-CoA may not be the factor that controls the flux from acetyl-CoA. The K_{m} for the second substrate, however, may be significant. AldDH has a K_{m} for NADH of 0.10mM, while PTA has a K_{m} for orthophosphate of 4.5mM. The intracellular concentrations of NADH and orthophosphate will influence the rate of flux from acetyl-CoA; however, these concentrations are not known from *G. thermoglucosidasius*. Despite this, one could assume that under fermentative growth on glucose the NADH concentration could be high due to increased rate of glycolysis under the micro-aerobic conditions. In addition, the concentration of orthophosphate may be low when high concentrations of glucose are available, because *G. thermoglucosidasius* uses the phosphotransferase system (PTS) to transport glucose into the cell. This involves the

transfer of the phosphoryl-group of phosphoenolpyruvate (PEP) onto the sugar once it has been transported into the cell to maintain the glucose concentration gradient across the cell membrane. The combination of high NADH concentration and low orthophosphate concentration could be a significant factor in controlling the flux from acetyl-CoA. Interestingly, pentose sugars such as xylose are metabolised by the pentose phosphate pathway which generates more NADPH instead of NADH. Therefore it could be conceived that the NADH concentrations are lower for xylose fermentation compared to glucose fermentation, resulting in a slower rate of acetyl-CoA to ethanol conversion, and this may be a significant factor at controlling the flux from acetyl-CoA.

The potential regulation of PTA was also studied in this Thesis. There is evidence in the literature that class II PTA enzymes are regulated by small metabolites, such as pyruvate, PEP, ADP, ATP and NADH (Alina Campos-Bermudez *et al.* 2010; Brinsmade and Escalante-Semerena 2007). However, there is little evidence that the class I PTA enzymes, which lack the N-terminal regulatory domain of the class II enzymes, are regulated. The results of Chapter 8 indicate that the *G. thermoglucosidasius* TM242 PTA, a member of the class I family, may be regulated by acetylation. The observed acetylation by acetyl-CoA was non-enzymatic and pH-dependent; an inactivation of PTA was observed at pH 8.0 but not at pH 6.5. This acetylation possibly occurs at a Lys residue at position 251 which has been shown to orientate the acetyl-phosphate substrate in the crystal structure of the highly similar PTA from *B. subtilis* (Xu *et al.* 2005). Acetylation at this position could disrupt substrate binding and hence inhibit the PTA enzyme. This would, however, require further studies to confirm this because another group suggests a Lys at position 250 is acetylated in the *B. subtilis* PTA (Kim *et al.* 2013).

The inactivation of PTA by acetylation under alkaline conditions hints at an acetate uptake role for PTA in *G. thermoglucosidasius*. Under alkaline conditions the PTA may be acetylated and inactivated, reducing the flux through to acetate. It is yet to be investigated whether this acetylation is reversible, but one could propose that as acetate is produced, either via the PTA-AK pathway or via an alternative, the pH level would drop leading to the activation of PTA which in combination with AK could catalyse the conversion of acetate to acetyl-CoA. The observation that a *pta* deletion prevents the re-assimilation of acetate (Chapter 7) supports this hypothesis. This also suggests that, although additional phosphotransacylases (PTACs) were discovered in *G. thermoglucosidasius* TM242, which could utilise acetyl-CoA,

the reverse flux through either phosphotranspropionylase (PTP) or phosphotransbutyrylase (PTB) was unable to compensate for the loss of PTA in terms of acetate re-assimilation.

Another common acetate uptake system in bacteria is via acetyl-CoA synthetase (ACS). *G. thermoglucosidasius* TM242 has four genes annotated as encoding ACS enzymes. Bioinformatics analysis revealed that the four ACS enzymes belong to the AMP-forming family, which, *in vivo*, catalyse the irreversible activation of acetate to acetyl-CoA, generating AMP and pyrophosphate. When recombinantly expressed, two of the proteins were found to be active ACS enzymes, while no activity was detected with ACS2. Further studies would be required to determine whether ACS2 is active with alternative substrates or whether it is a non-functional protein. Additional studies would also be necessary to determine whether *acs1* encodes a functional ACS enzyme. Despite two of the recombinant ACS enzymes being functional proteins, no ACS activity was detected from native *G. thermoglucosidasius* cell lysates suggesting that either the enzymes were not expressed or that they were inactive under the conditions measured. Initial studies suggested that *G. thermoglucosidasius* TM242 has an ACS regulatory system encoded by the *acuABC* operon, but the results were inconclusive and further studies would be required to confirm the true function of this operon.

The possibility of engineering a strain with reduced acetate production was investigated by disrupting the PTA-AK pathway via a gene-deletion strategy. The deletion of *pta* resulted in a strain that grew at a faster rate and potentially could produce ethanol at a faster rate than TM242. However, the Δ*pta* strain still produced acetate despite a reduction of 95% of PTA-specific activity (acetyl-CoA + orthophosphate → acetyl-phosphate + CoA). Through multiple combinations of phosphotransacylase (PTAC) deletions it was discovered that phosphotransbutyrylase (PTB) was responsible for the remaining PTA-specific activity in the Δ*pta* strain, suggesting that PTB could be involved in acetate production in this strain. However, complete disruption of PTA-specific activity, in either the Δ*pta*/Δ*ptb* or Δ*pta*/Δ*ptp*/Δ*ptb* strains, did not abolish acetate production (where *ptp* encoded a phosphotranspropionylase). Moreover, acetate yields increased when these two strains were grown on xylose. These results indicate that alternative pathway(s) for acetate production must be present in *G. thermoglucosidasius* TM242. Some possible alternatives were discussed in Chapter 7, including pyruvate oxidase and phosphoketolase, but these were ruled out based on bioinformatics results.

It has been reported that during exponential growth of *B. subtilis* the over-acidification of the medium due to the formation of pyruvate and acetate is prevented by conversion of pyruvate to uncharged acetoin under aerobic conditions, or uncharged 2,3-butanediol under anaerobic conditions (Speck and Freese 1973). Acetate can then be generated during the late exponential growth phase by the oxidative dissimilation of acetoin to acetaldehyde and subsequently to acetate (Lopez *et al.* 1975). In addition, acetoin can be oxidised to acetate via the butanediol cycle during the stationary phase (Speck and Freese 1973). The increased production of acetoin in the Δ*pta* strain suggests that *G. thermoglucosidasius* could produce acetate via a similar pathway to *B. subtilis* (Figure 7.22).

Acetate production via acetoin or 2,3-butanediol would not generate the important energy metabolite ATP. This may be critical to the survival and growth of *G. thermoglucosidasius* under micro-aerobic conditions because generation of ATP via the oxidative phosphorylation pathway would not be available due to the lack of the electron acceptor oxygen. PTA-specific activity was abolished in both the Δ*pta*/Δ*ptb* and Δ*pta*/Δ*ptp*/Δ*ptb* strains. Both these strains could ferment glucose, suggesting that the ATP yield from glycolysis (2 ATP molecules per molecule of glucose consumed) was sufficient for survival. However, they were both unable to fully utilise xylose suggesting that the ATP yield from the pentose phosphate pathway (1.67 ATP molecules per molecule of xylose consumed) was not sufficient to provide the energy requirement of micro-aerobic growth. This also suggests that the ATP yield is in a fine balance to support growth, and that the ATP generated from the PTA-AK pathway is necessary to support micro-aerobic growth of *G. thermoglucosidasius* TM242 on xylose.

In summary, the work conducted in this Thesis has demonstrated that acetate metabolism in *G. thermoglucosidasius* TM242 is necessary under micro-aerobic growth on xylose, and is controlled by a complex myriad of factors. The results suggest that the ATP generated in combination with acetate from the PTA-AK pathway is critical for micro-aerobic growth on pentose sugars such as xylose. This suggests that improvement to lignocellulosic bioethanol production cannot be made by the removal of the PTA-AK pathway, but perhaps could be made by the down regulation of its expression.

Chapter 10

References

Alina Campos-Bermudez, V., Pablo Bologna, F., Santiago Andreo, C. & Fabiana Drincovich, M., 2010. Functional dissection of *Escherichia coli* phosphotransacetylase structural domains and analysis of key compounds involved in activity regulation. *Febs Journal,* 277(8), pp. 1957-1966.

Alonso, D.M., Bond, J.Q. & Dumesic, J.A., 2010. Catalytic conversion of biomass to biofuels. *Green Chemistry,* 12(9), pp. 1493-1513.

Alvira, P., Tomas-Pejo, E., Ballesteros, M. & Negro, M.J., 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technology,* 101(13), pp. 4851-4861.

Argyros, D.A., Tripathi, S.A., Barrett, T.F., Rogers, S.R., Feinberg, L.F., Olson, D.G., Foden, J.M., Miller, B.B., Lynd, L.R., Hogsett, D.A. & Caiazza, N.C., 2011. High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. *Applied and Environmental Microbiology,* 77(23), pp. 8288-8294.

Atteia, A., van Lis, R., Gelius-Dietrich, G., Adrait, A., Garin, J., Joyard, J., Rolland, N. & Martin, W., 2006. Pyruvate formate-lyase and a novel route of eukaryotic ATP synthesis in *Chlamydomonas* mitochondria. *Journal of Biological Chemistry,* 281(15), pp. 9909-9918.

Averhoff, B., 2004. DNA transport and natural transformation in mesophilic and thermophilic bacteria. *Journal of Bioenergetics and Biomembranes,* 36(1), pp. 25-33.

Barnard, D., Casanueva, A., Tuffin, M. & Cowan, D., 2010. Extremophiles in biofuel synthesis. *Environmental Technology,* 31(8-9), pp. 871-888.

Bartosiak-Jentys, J., Hussein, A.H., Lewis, C.J. & Leak, D.J., 2013. Modular system for assessment of glycosyl hydrolase secretion in *Geobacillus thermoglucosidasius*. *Microbiology-Sgm,* 159, pp. 1267-1275.

Ben-Bassat, A., Lamed, R. & Zeikus, J.G., 1981. Ethanol production by thermophilic bacteria - Metabolic control of end product formation in *Thermoanaerobium brockii*. *Journal of Bacteriology,* 146(1), pp. 192-199.

Bentley, R., 2000. From 'reactive C-2 units' to acetyl coenzyme A: a long trail with an acetyl phosphate detour. *Trends in Biochemical Sciences,* 25(6), pp. 302-305.

Bettiga, M., Bengtsson, O., Hahn-Hagerdal, B. & Gorwa-Grauslund, M.F., 2009. Arabinose and xylose fermentation by recombinant *Saccharomyces cerevisiae* expressing a fungal pentose utilization pathway. *Microbial Cell Factories,* 8.

Bhalla, A., Bansal, N., Kumar, S., Bischoff, K.M. & Sani, R.K., 2013. Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresource Technology,* 128, pp. 751-759.

Bhandiwad, A., Shaw, A.J., Guss, A., Guseva, A., Bahl, H. & Lynd, L.R., 2014. Metabolic engineering of *Thermoanaerobacterium saccharolyticum* for n-butanol production. *Metabolic Engineering,* 21, pp. 17-25.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., ColladoVides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. & Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. *Science,* 277(5331), pp. 1453-1462.

Botting C.H. 2013. http://www.st-andrews.ac.uk/~bmsmspf/Website_html/

Bradford, M.M., 1976. Rapid and sensitive method for quantification of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry,* 72(1-2), pp. 248-254.

Brat, D., Boles, E. & Wiedemann, B., 2009. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology,* 75(8), pp. 2304-2311.

BRENDA enzyme database. http://www.brenda-enzymes.org/

Brinsmade, S.R. & Escalante-Semerena, J.C., 2004. The *eutD* gene of *Salmonella enterica* encodes a protein with phosphotransacetylase enzyme activity. *Journal of Bacteriology,* 186(6), pp. 1890-1892.

Brinsmade, S.R. & Escalante-Semerena, J.C., 2007. In vivo and in vitro analyses of single-mino acid variants of the *Salmonella enterica* phosphotransacetylase enzyme provide insights into the function of its N-terminal domain. *Journal of Biological Chemistry,* 282(17), pp. 12629-12640.

Brown, T.D.K., Jonesmortimer, M.C. & Kornberg, H.L., 1977. Enzymatic interconversion of acetate and acetyl-coenzyme-A in *Escherichia coli*. *Journal of General Microbiology,* 102(OCT), pp. 327-336.

Brown, T.D.K., Stormer, F.C. & Pereira, C.R.S., 1972. Studies of acetate kinase-phosphotransacetylase and butanediol-forming systems in *Aerobacter aerogenes*. *Journal of Bacteriology,* 112(3), pp. 1106-1111.

Burdette, D.S., Jung, S.H., Shen, G.J., Hollingsworth, R.I. & Zeikus, J.G., 2002. Physiological function of alcohol dehydrogenases and long-chain (C-30) fatty acids in alcohol tolerance of *Thermoanaerobacter ethanolicus*. *Applied and Environmental Microbiology,* 68(4), pp. 1914-1918.

Cava, F., Hidalgo, A. & Berenguer, J., 2009. Thermus thermophilus as biological model. *Extremophiles,* 13(2), pp. 213-231.

Cesar, C.E., Alvarez, L., Bricio, C., van Heerden, E., Littauer, D. & Berenguer, J., 2011. Unconventional lateral gene transfer in extreme thermophilic bacteria. *International Microbiology,* 14(4), pp. 187-199.

Chang, D.E., Shin, S., Rhee, J.S. & Pan, J.G., 1999. Acetate metabolism in a *pta* mutant of *Escherichia coli* W3110: Importance of maintaining acetyl coenzyme a flux for growth and survival. *Journal of Bacteriology,* 181(21), pp. 6656-6663.

Chang, T. & Yao, S., 2011. Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. *Applied Microbiology and Biotechnology,* 92(1), pp. 13-27.

Chen, R.Z., Yap, W., Postma, P.W. & Bailey, J.E., 1997. Comparative studies of *Escherichia coli* strains using different glucose uptake systems: Metabolism and energetics. *Biotechnology and Bioengineering,* 56(5), pp. 583-590.

Claassen, P.A.M., van Lier, J.B., Contreras, A.M.L., van Niel, E.W.J., Sijtsma, L., Stams, A.J.M., de Vries, S.S. & Weusthuis, R.A., 1999. Utilisation of biomass for the supply of energy carriers. *Applied Microbiology and Biotechnology,* 52(6), pp. 741-755.

Cooper, R.A., 1978. *Intermediary metabolism of monosaccharides by bacteria.* University Park Press: Baltimore, Md., USA.

Crabtree, H.G., 1929. Observations on the carbohydrate metabolism of tumours. *The Biochemical journal,* 23(3), pp. 536-45.

Cripps, R.E., Eley, K., Leak, D.J., Rudd, B., Taylor, M., Todd, M., Boakes, S., Martin, S. & Atkinson, T., 2009. Metabolic engineering of *Geobacillus thermoglucosidasius* for high yield ethanol production. *Metabolic Engineering,* 11(6), pp. 398-408.

Culha, M., Adiguzel, A., Yazici, M.M., Kahraman, M., Sahin, F. & Gulluce, M., 2008. Characterization of thermophilic bacteria using surface-enhanced raman scattering. *Applied Spectroscopy,* 62(11), pp. 1226-1232.

Dahms, A.S., 1974. 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. *Biochemical and Biophysical Research Communications,* 60(4), pp. 1433-1439.

De Mey, M., De Maeseneire, S., Soetaert, W. & Vandamme, E., 2007. Minimizing acetate formation in *E. coli* fermentations. *Journal of Industrial Microbiology & Biotechnology,* 34(11), pp. 689-700.

Demain, A.L., Newcomb, M. & Wu, J.H.D., 2005. Cellulase, clostridia, and ethanol. *Microbiology and Molecular Biology Reviews,* 69(1), pp. 124-154.

Deutscher, J., 2008. The mechanisms of carbon catabolite repression in bacteria. *Current Opinion in Microbiology,* 11(2), pp. 87-93.

Dien, B.S., Cotta, M.A. & Jeffries, T.W., 2003. Bacteria engineered for fuel ethanol production: current status. *Applied Microbiology and Biotechnology,* 63(3), pp. 258-266.

Dimou, M., Venieraki, A., Liakopoulos, G. & Katinakis, P., 2011. Cloning, characterization and transcriptional analysis of two phosphate acetyltransferase isoforms from *Azotobacter vinelandii*. *Molecular Biology Reports,* 38(6), pp. 3653-3663.

Eglinton, J.M., Heinrich, A.J., Pollnitz, A.P., Langridge, P., Henschke, P.A. & Lopes, M.D., 2002. Decreasing acetic acid accumulation by a glycerol overproducing strain of *Saccharomyces cerevisiae* by deleting the *ALD6* aldehyde dehydrogenase gene. *Yeast,* 19(4), pp. 295-301.

El-Mansi, M., 2004. Flux to acetate and lactate excretions in industrial fermentations: physiological and biochemical implications. *Journal of Industrial Microbiology & Biotechnology,* 31(7), pp. 295-300.

Elkins, J.G., Raman, B. & Keller, M., 2010. Engineered microbial systems for enhanced conversion of lignocellulosic biomass. *Current Opinion in Biotechnology,* 21(5), pp. 657-662.

Ellman, G.L., 1959. Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics,* 82(1), pp. 70-77.

Elmansi, E.M.T. & Holms, W.H., 1989. Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures. *Journal of General Microbiology,* 135, pp. 2875-2883.

Entner, N. & Doudoroff, M., 1952. Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. *Journal of Biological Chemistry,* 196(2), pp. 853-862.

Extance, J.P., 2012. Bioethanol production: Characterisation of a bifunctional alcohol dehydrogenase from *Geobacillus thermoglucosidasius*. Ph.D thesis. Thesis, University of Bath.

Extance, J., Crennell, S.J., Eley, K., Cripps, R., Hough, D.W. & Danson, M.J., 2013. Structure of a bifunctional alcohol dehydrogenase involved in bioethanol generation in *Geobacillus thermoglucosidasius*. *Acta crystallographica section D - Biological crystallography,* 69, pp. 2104-2115.

Finkel, T., Deng, C.-X. & Mostoslavsky, R., 2009. Recent progress in the biology and physiology of sirtuins. *Nature,* 460(7255), pp. 587-591.

Fong, J.C.N., Svenson, C.J., Nakasugi, K., Leong, C.T.C., Bowman, J.P., Chen, B., Glenn, D.R., Neilan, B.A. & Rogers, P.L., 2006. Isolation and characterization of two novel ethanol-tolerant facultative-anaerobic thermophilic bacteria strains from waste compost. *Extremophiles,* 10(5), pp. 363-372.

Fujita, Y., 2009. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. *Bioscience Biotechnology and Biochemistry,* 73(2), pp. 245-259.

Gardner, J.G., Grundy, F.J., Henkin, T.M. & Escalante-Semerena, J.C., 2006. Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in *Bacillus subtilis*. *Journal of Bacteriology,* 188(15), pp. 5460-5468.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. & Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server. Walker, John M. ed.: The Proteomics Protocols Handbook, *Humana Press*, pp. 571-607.

Gerosa, L. & Sauer, U., 2011. Regulation and control of metabolic fluxes in microbes. *Current Opinion in Biotechnology,* 22(4), pp. 566-575.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., III & Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods,* 6(5), pp. 343-345.

Green, E.M., Boynton, Z.L., Harris, L.M., Rudolph, F.B., Papoutsakis, E.T. & Bennett, G.N., 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology-Uk,* 142, pp. 2079-2086.

Grundy, F.J., Turinsky, A.J. & Henkin, T.M., 1994. Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilisation genes by CcpA. *Journal of Bacteriology,* 176(15), pp. 4527-4533.

Grundy, F.J., Waters, D.A., Allen, S.H.G. & Henkin, T.M., 1993a. Regulation of the *Bacillus subtilis* acetate kinase gene by CcpA. *Journal of Bacteriology,* 175(22), pp. 7348-7355.

Grundy, F.J., Waters, D.A., Takova, T.Y. & Henkin, T.M., 1993b. Identifiction of genes involved in utilization of acetate and acetoin in *Bacillus subtilis*. *Molecular Microbiology,* 10(2), pp. 259-271.

Haber, W., 2007. Energy, food, and land - The ecological traps of humankind. *Environmental Science and Pollution Research,* 14(6), pp. 359-365.

Hild, H.M., Stuckey, D.C. & Leak, D.J., 2003. Effect of nutrient limitation on product formation during continuous fermentation of xylose with *Thermoanaerobacter ethanolicus* JW200 Fe(7). *Applied Microbiology and Biotechnology,* 60(6), pp. 679-686.

Hu, L.I., Lima, B.P. & Wolfe, A.J., 2010. Bacterial protein acetylation: the dawning of a new age. *Molecular Microbiology,* 77(1), pp. 15-21.

Huang, M., Oppermann-Sanio, F.B. & Steinbuchel, A., 1999. Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. *Journal of Bacteriology,* 181(12), pp. 3837-3841.

Ingram-Smith, C., Gorrell, A., Lawrence, S.H., Iyer, P., Smith, K. & Ferry, J.G., 2005. Characterization of the acetate binding pocket in the *Methanosarcina thermophila* acetate kinase. *Journal of Bacteriology,* 187(7), pp. 2386-2394.

Ingram-Smith, C., Martin, S.R. & Smith, K.S., 2006. Acetate kinase: not just a bacterial enzyme. *Trends in Microbiology,* 14(6), pp. 249-253.

Jones, J.D. & O'Connor, C.D., 2011. Protein acetylation in prokaryotes. *Proteomics,* 11(15), pp. 3012-3022.

Jordan, D.B., Bowman, M.J., Braker, J.D., Dien, B.S., Hector, R.E., Lee, C.C., Mertens, J.A. & Wagschal, K., 2012. Plant cell walls to ethanol. *Biochemical Journal,* 442, pp. 241-252.

Kahane, I. & Muhlrad, A., 1979. Purification and properties of acetate kinase from *Acholeplasma laidlawii*. *Journal of Bacteriology,* 137(2), pp. 764-772.

Keller, M.W., Schut, G.J., Lipscomb, G.L., Menon, A.L., Iwuchukwu, I.J., Leuko, T.T., Thorgersen, M.P., Nixon, W.J., Hawkins, A.S., Kelly, R.M. & Adams, M.W.W., 2013. Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proceedings of the National Academy of Sciences of the United States of America,* 110(15), pp. 5840-5845.

Kerr, R.A., 2007. Global warming is changing the world. *Science,* 316(5822), pp. 188-190.

Kim, D., Yu, B.J., Kim, J.A., Lee, Y.-J., Choi, S.-G., Kang, S. & Pan, J.-G., 2013. The acetylproteome of Gram-positive model bacterium *Bacillus subtilis*. *Proteomics,* 13(10-11), pp. 1726-1736.

Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K. & Mardis, E.R., 2013. The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell,* 155(1), pp. 27-38.

Kuit, W., Minton, N.P., Lopez-Contreras, A.M. & Eggink, G., 2012. Disruption of the acetate kinase (*ack*) gene of *Clostridium acetobutylicum* results in delayed acetate production. *Applied Microbiology and Biotechnology,* 94(3), pp. 729-741.

Kumar, S., Singh, S.P., Mishra, I.M. & Adhikari, D.K., 2009. Recent advances in production of bioethanol from lignocellulosic biomass. *Chemical Engineering & Technology,* 32(4), pp. 517-526.

Lacis, L.S. & Lawford, H.G., 1991. *Thermoanaerobacter ethanolicus* growth and product yield from elevated levels of xylose or glucose in continuous cultures. *Applied and Environmental Microbiology,* 57(2), pp. 579-585.

Latimer, M.T. & Ferry, J.G., 1993. Cloning, sequence analysis, and hyperexpression of the genes encoding phosphotransacetylase and acetate kinase from *Methanosarcina thermophila*. *Journal of Bacteriology,* 175(21), pp. 6822-6829.

Lee, D.-W., Kim, D., Lee, Y.-J., Kim, J.-A., Choi, J.Y., Kang, S. & Pan, J.-G., 2013. Proteomic analysis of acetylation in thermophilic *Geobacillus kaustophilus*. *Proteomics,* 13(15), pp. 2278-2282.

Levine, S.M., Ardeshir, F. & Ames, G.F.L., 1980. Isolation and characteriastion of acetate kinase and phosphotransacetylase mutants of *Echerichia coli* and *Salmonella typhimurium*. *Journal of Bacteriology,* 143(2), pp. 1081-1085.

Liao, H.H. & Kanikula, A.M., 1990. Increased efficiency of transformation of *Bacillus stearothermophilus* by a plasmid carrying a thermostable kanamycin resistance marker. *Current Microbiology,* 21(5), pp. 301-306.

Lin, L., Song, H., Ji, Y., He, Z., Pu, Y., Zhou, J. & Xu, J., 2010. Ultrasound-mediated DNA transformation in thermophilic Gram-positive anaerobes. *Plos One,* 5(9).

Lin, L. & Xu, J., 2013. Dissecting and engineering metabolic and regulatory networks of thermophilic bacteria for biofuel production. *Biotechnology Advances,* 31(6), pp. 827-837.

Lin, P.P., Rabe, K.S., Takasumi, J.L., Kadisch, M., Arnold, F.H. & Liao, J.C., 2014. Isobutanol production at elevated temperatures in thermophilic *Geobacillus thermoglucosidasius*. *Metabolic Engineering,* 24, pp. 1-8.

Liu, Y., Leal, N.A., Sampson, E.M., Johnson, C.L.V., Havemann, G.D. & Bobik, T.A., 2007. PduL is an evolutionarily distinct phosphotransacylase involved in B-12-dependent 1,2-propanediol degradation by *Salmonella enterica* serovar typhimurium LT2. *Journal of Bacteriology,* 189(5), pp. 1589-1596.

Lopez, J.M., Thoms, B. & Rehbein, H., 1975. Acetoin degradation in *Bacillus subtilis* by direct oxidative cleavage. *European Journal of Biochemistry,* 57(2), pp. 425-430.

Lynd, L.R., Laser, M.S., Bransby, D., Dale, B.E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillar, J.D., Sheehan, J. & Wyman, C.E., 2008. How biotech can transform biofuels. *Nature Biotechnology,* 26(2), pp. 169-172.

MacLean, D., Jones, J.D.G. & Studholme, D.J., 2009. Application of 'next-generation' sequencing technologies to microbial genetics. *Nature Reviews Microbiology,* 7(4), pp. 287-296.

Majewski, R.A. & Domach, M.M., 1990. Simple constrained optimization view of acetate overflow in *Escherichia coli*. *Biotechnology and Bioengineering,* 35(7), pp. 732-738.

McCleary, W.R., Stock, J.B. & Ninfa, A.J., 1993. Is acetyl phosphate a global signal in *Escherichia coli*? *Journal of Bacteriology,* 175(10), pp. 2793-2798.

McKenzie, T., Hoshino, T., Tanaka, T. & Sueoka, N., 1986. The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. *Plasmid,* 15(2), pp. 93-103.

Moir-Blais, T.R., Grundy, F.J. & Henkin, T.M., 2001. Transcriptional activation of the *Bacillus subtilis ackA* promoter requires sequences upstream of the CcpA binding site. *Journal of Bacteriology,* 183(7), pp. 2389-2393.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M. & Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology,* 96(6), pp. 673-686.

Nakajima, H., Suzuki, K. & Imahori, K., 1978. Purification and properties of acetate kinase from *Bacillus stearothermophilus*. *Journal of Biochemistry,* 84(1), pp. 193-203.

Nazina, T.N., Tourova, T.P., Poltaraus, A.B., Novikova, E.V., Grigoryan, A.A., Ivanova, A.E., Lysenko, A.M., Petrunyaka, V.V., Osipov, G.A., Belyaev, S.S. & Ivanov, M.V., 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp nov and *Geobacillus uzenensis* sp nov from petroleum reservoirs and transfer of *Bacillus stearothermophilus Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. thermocatenulatus*, *G. thermoleovorans*, *G. kaustophilus*, *G. thermoglucosidasius* and *G. thermodenitrificans*. *International Journal of Systematic and Evolutionary Microbiology,* 51, pp. 433-446.

Olson, D.G. & Lynd, L.R., 2012a. Computational design and characterization of a temperature-sensitive plasmid replicon for gram positive thermophiles. *Journal of Biological Engineering,* 6(1).

Olson, D.G. & Lynd, L.R., 2012b. Transformation of *Clostridium thermocellum* by electroporation. *Cellulases,* 510, pp. 317-330.

Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B. & Keasling, J.D., 2012. Microbial engineering for the production of advanced biofuels. *Nature,* 488(7411), pp. 320-328.

Phue, J.-N., Lee, S.J., Kaufman, J.B., Negrete, A. & Shiloach, J., 2010. Acetate accumulation through alternative metabolic pathways in *ackA*$^{-}$ *pta*$^{-}$ *poxB*$^{-}$ triple mutant in *E. coli* B (BL21). *Biotechnology Letters,* 32(12), pp. 1897-1903.

Ponce, E., 1999. Effect of growth rate reduction and genetic modifications on acetate accumulation and biomass yields in *Escherichia coli*. *Journal of Bioscience and Bioengineering,* 87(6), pp. 775-780.

Presecan-Siedel, E., Galinier, A., Longin, R., Deutscher, J., Danchin, A., Glaser, P. & Martin-Verstraete, I., 1999. Catabolite regulation of the *pta* gene as part of carbon flow pathways in *Bacillus subtilis*. *Journal of Bacteriology,* 181(22), pp. 6889-6897.

Promega. Wizard Plus SV Minipreps DNA Purification System.

Promega. Wizard SV Gel and PCR Clean-Up System.

Rado, T.A. & Hoch, J.A., 1973. Phosphotransacetylase from *Bacillus subtilis* - Purification and physiological studies. *Biochimica Et Biophysica Acta,* 321(1), pp. 114-125.

Renewable Fuel Association (RFA). (2014). 2014 Ethanol industry outlook [online]. Retrieved 13th August, 2014, from http://www.ethanolrfa.org/page/-/rfa-association-site/Resource%20Center/2014%20Ethanol%20Industry%20Outlook.pdf?nocdn=1

Richardson, E.J. & Watson, M., 2013. The automatic annotation of bacterial genomes. *Briefings in Bioinformatics,* 14(1), pp. 1-12.

Rude, M.A. & Schirmer, A., 2009. New microbial fuels: a biotech perspective. *Current Opinion in Microbiology,* 12(3), pp. 274-281.

San Martin, R., Bushell, D., Leak, D.J. & Hartley, B.S., 1992. Development of a synthetic medium for continuous anaerobic growth and ethanol production with a lactate dehydrogenase mutant of *Bacillus stearothermophilus*. *Journal of general microbiology,* 138(5), pp. 987-96.

Sanchez, O.J. & Cardona, C.A., 2008. Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresource Technology,* 99(13), pp. 5270-5295.

Semenza, G., 2001. The identification of 'active acetate' as acetyl-CoA. *Febs Letters,* 509(3), pp. 343-344.

Shaw, A.J., Podkaminer, K.K., Desai, S.G., Bardsley, J.S., Rogers, S.R., Thorne, P.G., Hogsett, D.A. & Lynd, L.R., 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proceedings of the National Academy of Sciences of the United States of America,* 105(37), pp. 13769-13774.

Shin, B.S., Choi, S.K. & Park, S.H., 1999. Regulation of the *Bacillus subtilis* phosphotransacetylase gene. *Journal of Biochemistry,* 126(2), pp. 333-339.

Silhavy, T.J., Kahne, D. & Walker, S., 2010. The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology,* 2(5).

Sillers, R., Chow, A., Tracy, B. & Papoutsakis, E.T., 2008. Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. *Metabolic Engineering,* 10(6), pp. 321-332.

Sommer, P., Georgieva, T. & Ahring, B.K., 2004. Potential for using thermophilic anaerobic bacteria for bioethanol production from hemicellulose. *Biochemical Society Transactions,* 32, pp. 283-289.

Soppa, J., 2010. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea-an International Microbiological Journal*.

Speck, E.L. & Freese, E., 1973. Control of metabolite secretion in *Bacillus subtilis*. *Journal of General Microbiology,* 78(OCT), pp. 261-275.

Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D. & Escalante-Semerena, J.C., 2002. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science,* 298(5602), pp. 2390-2392.

Starai, V.J. & Escalante-Semerena, J.C., 2004a. Acetyl-coenzyme A synthetase (AMP forming). *Cellular and Molecular Life Sciences,* 61(16), pp. 2020-2030.

Starai, V.J. & Escalante-Semerena, J.C., 2004b. Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. *Journal of Molecular Biology,* 340(5), pp. 1005-1012.

Starai, V.J., Garrity, J. & Escalante-Semerena, J.C., 2005. Acetate excretion during growth of *Salmonella enterica* on ethanolamine requires phosphotransacetylase (EutD) activity, and acetate recapture requires acetyl-CoA synthetase (Acs) and phosphotransacetylase (Pta) activities. *Microbiology-Sgm,* 151, pp. 3793-3801.

Stephanopoulos, G., 2007. Challenges in engineering microbes for biofuels production. *Science,* 315(5813), pp. 801-804.

Stothard, P. & Wishart, D.S., 2006. Automated bacterial genome analysis and annotation. *Current Opinion in Microbiology,* 9(5), pp. 505-510.

Sun, Y. & Cheng, J.Y., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology,* 83(1), pp. 1-11.

Suzuki, H., Murakami, A. & Yoshida, K.-i., 2012. Counterselection System for *Geobacillus kaustophilus* HTA426 through Disruption of *pyrF* and *pyrR*. *Applied and Environmental Microbiology,* 78(20), pp. 7376-7383.

Taylor, M.P., Eley, K.L., Martin, S., Tuffin, M.I., Burton, S.G. & Cowan, D.A., 2009. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. *Trends in Biotechnology,* 27(7), pp. 398-405.

Taylor, M.P., Esteban, C.D. & Leak, D.J., 2008. Development of a versatile shuttle vector for gene expression in *Geobacillus* spp. *Plasmid,* 60(1), pp. 45-52.

Tenenbaum, D.J., 2008. Food vs. fuel: Diversion of crops could cause more hunger. *Environmental Health Perspectives,* 116(6), pp. A254-A257.

Thompson, D.K. & Chen, J.S., 1990. Purification and properties of an acetoacetyl coenzyme A reacting phosphotransbutyrylase from *Clostridium beijerinckii ("Clostridium butyl*icum") NRRL B593. *Applied and Environmental Microbiology,* 56(3), pp. 607-613.

TMO Renewables Ltd. 40 Alan Turing Road, Surrey Research Park, Guildford. GU2 7YF.

Tripathi, S.A., Olson, D.G., Argyros, D.A., Miller, B.B., Barrett, T.F., Murphy, D.M., McCool, J.D., Warner, A.K., Rajgarhia, V.B., Lynd, L.R., Hogsett, D.A. & Caiazza, N.C., 2010. Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. *Applied and Environmental Microbiology,* 76(19), pp. 6591-6599.

van den Berg, M.A., de Jong Gubbels, P., Kortland, C.J., van Dijken, J.P., Pronk, J.T. & Steensma, H.Y., 1996. The two acetyl-coenzyme A synthetases of *Saccharomyces cerevisiae* differ with respect to kinetic properties and transcriptional regulation. *Journal of Biological Chemistry,* 271(46), pp. 28953-28959.

Van Zyl, L.J., Taylor, M.P., Eley, K., Tuffin, M. & Cowan, D.A., 2014. Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host *Geobacillus thermoglucosidasius*. *Applied Microbiology and Biotechnology,* 98(3), pp. 1247-1259.

Verdin, E. & Ott, M., 2013. Acetylphosphate: A novel link between lysine acetylation and intermediary metabolism in bacteria. *Molecular Cell,* 51(2), pp. 132-134.

Wang, J., Huang, J., Shi, J., Xu, Q., Xie, X. & Chen, N., 2013. Fermentation characterization of an L-tryptophan producing *Escherichia coli* strain with inactivated phosphotransacetylase. *Annals of Microbiology,* 63(4), pp. 1219-1224.

Wang, M., Fu, J., Zhang, X. & Chen, T., 2012. Metabolic engineering of *Bacillus subtilis* for enhanced production of acetoin. *Biotechnology Letters,* 34(10), pp. 1877-1885.

Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.-B., Zeng, R., Xiong, Y., Guan, K.-L., Zhao, S. & Zhao, G.-P., 2010. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science,* 327(5968), pp. 1004-1007.

Wanner, B.L., 1993. Gene regulation by phosphate in enteric bacteria. *Journal of Cellular Biochemistry,* 51(1), pp. 47-54.

Weimberg, R., 1961. Pentose oxidation by *Pseudomonas fragi*. *Journal of Biological Chemistry,* 236(3), pp. 629-635.

Weinert, B.T., Iesmantavicius, V., Wagner, S.A., Scholz, C., Gummesson, B., Beli, P., Nystrom, T. & Choudhary, C., 2013. Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Molecular Cell,* 51(2), pp. 265-272.

Wiesenborn, D.P., Rudolph, F.B. & Papoutsakis, E.T., 1989. Phosphotransbutyrylase from *Clostridium acetobutylicum* ATCC 824 and its role in acidogenesis. *Applied and Environmental Microbiology,* 55(2), pp. 317-322.

Winzer, K., Lorenz, K. & Durre, P., 1997. Acetate kinase from *Clostridium acetobutylicum*: a highly specific enzyme that is actively transcribed during acidogenesis and solventogenesis. *Microbiology-Uk,* 143, pp. 3279-3286.

Wolfe, A.J., 2005. The acetate switch. *Microbiology and Molecular Biology Reviews,* 69(1), pp. 12-50.

Wolfe, A.J., 2008. Quorum sensing "flips" the acetate switch. *Journal of Bacteriology,* 190(17), pp. 5735-5737.

Wolfe, A.J., 2010. Physiologically relevant small phosphodonors link metabolism to signal transduction. *Current Opinion in Microbiology,* 13(2), pp. 204-209.

Wu, S., Liu, B. & Zhang, X., 2006. Characterization of a recombinant thermostable xylanase from deep-sea thermophilic *Geobacillus* sp MT-1 in East Pacific. *Applied Microbiology and Biotechnology,* 72(6), pp. 1210-1216.

Xu, Q.S., Jancarik, J., Lou, Y., Kuznetsova, K., Yakunin, A.F., Yokota, H., Adams, P., Kim, R. & Kim, S.-H., 2005. Crystal structures of a phosphotransacetylase from *Bacillus subtilis* and its complex with acetyl phosphate. *Journal of Structural and Functional Genomics,* 6(4), pp. 269-279.

Yoshimura, F., 1978. Purification and characterisation of acetate kinase from *Veillonella alcalescens* ATCC 17748. *Archives of Biochemistry and Biophysics,* 189(2), pp. 424-432.

Yu, B.J., Kim, J.A., Moon, J.H., Ryu, S.E. & Pan, J.-G., 2008. The diversity of lysine-acetylated proteins in *Escherichia coli*. *Journal of Microbiology and Biotechnology,* 18(9), pp. 1529-1536.

Zaldivar, J., Nielsen, J. & Olsson, L., 2001. Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Applied Microbiology and Biotechnology,* 56(1-2), pp. 17-34.

Zalieckas, J.M., Wray, L.V. & Fisher, S.H., 1998. Expression of the *Bacillus subtilis acsA* gene: Position and sequence context affect *cre*-mediated carbon catabolite repression. *Journal of Bacteriology,* 180(24), pp. 6649-6654.

Zeikus, J.G., Ben-Bassat, A., Ng, T.K. & Lamed, R.J., 1981. Thermophilic ethanol fermentations. *Basic life sciences,* 18, pp. 441-61.

Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Li, Y., Shi, J., An, W., Hancock, S.M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y. & Guan, K.-L., 2010. Regulation of cellular metabolism by protein lysine acetylation. *Science,* 327(5968), pp. 1000-1004.

Zhu, Y., Liu, X.G. & Yang, S.T., 2005. Construction and characterization of *pta* gene-deleted mutant of *Clostridium tyrobutyricum* for enhanced butyric acid fermentation. *Biotechnology and Bioengineering,* 90(2), pp. 154-166.

Appendix

Full-length phosphotransacetylase

Coding DNA sequence:

```
ATGGCAATGCTTAACTTAGTTGCGTTTCTATATAATAAGAGTATCCTAGTATATAGGGGCATTTCAAACATTAGGG
AGGATTCGTTCGTGAGCAGTGATTTATTTTCGACATTAAAAGAAAAAATAGCGGGAAAACAACGGAAAATCGTGTT
TCCGGAAGGGCTTGATGAGCGTATTTTAACAGCGGTAAGCCGTCTGGCGAACGAGCAAATCGTCACGCCGATTGTC
ATTGGCAATGAAGAAGCGGTTAAGCAAAAAGCAAGCGAGCTTGGGCTGACGCTTCCGAATGTCGAAATCATTGATC
CGCATCAGTACGGGGAAATGGACAAGCTTGTTGCGGCATTTGTCGAACGCCGCAAAGGGAAAGTGACGGAAGAAGC
GGCGCGGAAGCTGCTTCTTGACGAAAATTATTTTGGCACCATGCTTGTGTACATGGATAAGGCGCATGGGCTTGTC
AGCGGCGCGGCGCATTCGACGGCTGATACGGTGCGGCCTGCGTTGCAAATTATAAAAACGAAACAAGGCGTCCGCA
AAACGTCAGGAGTATTCATTATGGTGCGCGGTGATGAAAAGTACGTGTTTGCCGATTGCGCGATCAACATTGCCCC
GGACAGCCAAGATTTGGCGGAAATCGCTGTCGAAAGCGCCAACACGGCAAAAATGTTCGACATTGAGCCGCGCGTG
GCGATGTTGAGCTTTTCGACAAAAGGATCAGCGAAATCGCCAGAAACGGAAAAAGTCGTCGAAGCGGTGCGGCTTG
CGAAAGAAATGGCGCCTGACTTAGTGCTGGACGGTGAGTTTCAGTTCGACGCGGCGTTTGTTCCGTCTGTCGCGAA
AAAGAAAGCGCCAGATTCCGTCATTCAAGGAGACGCGAACGTATTTATTTTCCCAAGCCTTGAAGCGGGAAATATC
GGCTATAAAATCGCCCAGCGTCTCGGCAACTTTGAAGCGGTCGGCCCGATTTTGCAAGGACTCAATAAGCCTGTGA
ACGACCTGTCACGCGGTTGCAATGCGGAAGATGTGTACAAGCTGACGCTTATAACTGCGGCGCAATCGCTATAA
```

The GTG codon that was identified as a second translation initiator and mutated to GTT in Chapter 4 is highlighted above.

Peptide sequence:

```
MAMLNLVAFLYNKSILVYRGISNIREDSFVSSDLFSTLKEKIAGKQRKIVFPEGLDERILTAVSRLANEQIVTPIV
IGNEEAVKQKASELGLTLPNVEIIDPHQYGEMDKLVAAFVERRKGKVTEEAARKLLLDENYFGTMLVYMDKAHGLV
SGAAHSTADTVRPALQIIKTKQGVRKTSGVFIMVRGDEKYVFADCAINIAPDSQDLAEIAVESANTAKMFDIEPRV
AMLSFSTKGSAKSPETEKVVEAVRLAKEMAPDLVLDGEFQFDAAFVPSVAKKKAPDSVIQGDANVFIFPSLEAGNI
GYKIAQRLGNFEAVGPILQGLNKPVNDLSRGCNAEDVYKLTLITAAQSL-
```

Truncated phosphotransacetylase

Coding DNA sequence:

```
GTGAGCAGTGATTTATTTTCGACATTAAAAGAAAAAATAGCGGGAAAACAACGGAAAATCGTGTTTCCGGAAGGGC
TTGATGAGCGTATTTTAACAGCGGTAAGCCGTCTGGCGAACGAGCAAATCGTCACGCCGATTGTCATTGGCAATGA
AGAAGCGGTTAAGCAAAAAGCAAGCGAGCTTGGGCTGACGCTTCCGAATGTCGAAATCATTGATCCGCATCAGTAC
GGGGAAATGGACAAGCTTGTTGCGGCATTTGTCGAACGCCGCAAAGGGAAAGTGACGGAAGAAGCGGCGCGGAAGC
TGCTTCTTGACGAAAATTATTTTGGCACCATGCTTGTGTACATGGATAAGGCGCATGGGCTTGTCAGCGGCGCGGC
GCATTCGACGGCTGATACGGTGCGGCCTGCGTTGCAAATTATAAAAACGAAACAAGGCGTCCGCAAAACGTCAGGA
GTATTCATTATGGTGCGCGGTGATGAAAAGTACGTGTTTGCCGATTGCGCGATCAACATTGCCCCGGACAGCCAAG
ATTTGGCGGAAATCGCTGTCGAAAGCGCCAACACGGCAAAAATGTTCGACATTGAGCCGCGCGTGGCGATGTTGAG
CTTTTCGACAAAAGGATCAGCGAAATCGCCAGAAACGGAAAAAGTCGTCGAAGCGGTGCGGCTTGCGAAAGAAATG
GCGCCTGACTTAGTGCTGGACGGTGAGTTTCAGTTCGACGCGGCGTTTGTTCCGTCTGTCGCGAAAAAGAAAGCGC
CAGATTCCGTCATTCAAGGAGACGCGAACGTATTTATTTTCCCAAGCCTTGAAGCGGGAAATATCGGCTATAAAAT
CGCCCAGCGTCTCGGCAACTTTGAAGCGGTCGGCCCGATTTTGCAAGGACTCAATAAGCCTGTGAACGACCTGTCA
CGCGGTTGCAATGCGGAAGATGTGTACAAGCTGACGCTTATAACTGCGGCGCAATCGCTATAA
```

Peptide sequence:

```
VSSDLFSTLKEKIAGKQRKIVFPEGLDERILTAVSRLANEQIVTPIVIGNEEAVKQKASELGLTLPNVEIIDPHQY
GEMDKLVAAFVERRKGKVTEEAARKLLLDENYFGTMLVYMDKAHGLVSGAAHSTADTVRPALQIIKTKQGVRKTSG
VFIMVRGDEKYVFADCAINIAPDSQDLAEIAVESANTAKMFDIEPRVAMLSFSTKGSAKSPETEKVVEAVRLAKEM
APDLVLDGEFQFDAAFVPSVAKKKAPDSVIQGDANVFIFPSLEAGNIGYKIAQRLGNFEAVGPILQGLNKPVNDLS
RGCNAEDVYKLTLITAAQSL-
```

Phosphotranspropionylase

Coding DNA sequence:

ATGAGCAATGAACTGCAAGACAAAATTTTCAAAGCAGTAATGAAAAGAATACAGGAAATAGAAGAAGGTCCATATG
TGCCAATCGGCGTTTCTAACCGGCATGTTCATTTGTCCCAAGAAGACTTAGAAAGTCTTTTCGGCAAAAACTACCG
CTTGACGAAATTAAGAGATTTAAAACAGCCGGGACAATTTGCTGCAAATGAAACCGTTACTCTCAAAGGGCCTAAA
GGAGAAATCCAACATGTAAGGATTTTAGGGCCGATAAGAAGTAAAACGCAAGTGGAAATATCGATTACAGATAGTT
TTCAGTTAGGGATTTCTGTACCAGTAAGAGAATCAGGAAATATTGCGGGCACACCAGGAATCATCATTCAAGGCCC
AAAAGGCATTGTGAAACTAAAAGAAGGTGTGATTGCGGCACTAAGGCACATACATGTTCCGCCAGAATTTGCAGCG
AAATTTCAATTAAAAGATAAGGACATGGTAGATGTTGAAGTTGGCAAAGAAAGAAAAACGATTTTTAAAAATGTTC
TTATTCGAGTTTCAAAAAATTACGTTTTAGAAATGCATTTAGATACAGATGAGGCGAATGCTGCAGGATTGAGAAA
CGGAGATATAGGAAAAATAGTGAAATGA

Peptide sequence:

MSNELQDKIFKAVMKRIQEIEEGPYVPIGVSNRHVHLSQEDLESLFGKNYRLTKLRDLKQPGQFAANETVTLKGPK
GEIQHVRILGPIRSKTQVEISITDSFQLGISVPVRESGNIAGTPGIIIQGPKGIVKLKEGVIAALRHIHVPPEFAA
KFQLKDKDMVDVEVGKERKTIFKNVLIRVSKNYVLEMHLDTDEANAAGLRNGDIGKIVK-

Phosphotransbutyrylase

Coding DNA sequence:

ATGAAGAGGTTGGTGCATAAGATGAAGCTGCAATCGTTGATCGATCGCACAACGAAATCCCAAAATATGACGGTAG
CTGTTGCGGCAGCGGAGGATGAAGAAGTCATCGATGCGGTGGCAATGGCGCTTGATCGCCATTTCGGGAAATTTGT
GTTGTATGGAGATAGGGAACAAATTTTGCAACTGTTAAAGCAAAAAAATTATGCAAATTCGGACAACATACAAATC
GTCCACGCCAATTCTGCCGTACAAGCGGCTGAGCTTGCGGTCAGAGCGGTTCATTCAAATGAAGCGGGCGTGCTGA
TGAAAGGAAATGTCCCGACCGCTGCGCTTTTAAAAGCGGTGTTAAACAAAGAATATGGTTTGCGCATAGGGAAAGT
GCTTTCCCATGTCGCGGTATTTGAAATCCCCGGCTACGACCGGCTCATCATTGTGACAGATGCCGCGATGAACGTT
TCACCAGATTTAGAGCAAAAAGCGCAAATTATTGCGAATGCGGTTGCGGTGGCCCATGCGATTGGCATTGAAACGC
CAAAAGTGGCACCGCTTGCGGCGGTGGAAATCGTGAACCCATCCATGCCGGCAACTGTCGATGCCGCGGCATTGGC
GATGATGCAAAAGCGCGGGCAAATAAAGGGATGTGTTATTGACGGCCCGCTCGCGCTTGATAACGCTGTTTCGCCC
AAGGCAGCCAAACATAAGCGGATTGATAGCGAAGTCGCAGGAAAAGCAGATATTTTGCTCGTTCCGGACATTGAAG
CGGGGAACATGCTGTATAAATCGCTCGTTTATTTTGCAAACGCAAAAGTCGGTGCAATCATCGCTGGGGCAAAAGC
GCCGATTGTCCTTACGTCGCGCGCCGATTCCGCAGAAAGCAAGCTTTACTCATTAGCCCTTGCTGTTTGTTCCGCA
TCGAAATAA

Peptide sequence:

MKRLVHKMKLQSLIDRTTKSQNMTVAVAAAEDEEVIDAVAMALDRHFGKFVLYGDREQILQLLKQKNYANSDNIQI
VHANSAVQAAELAVRAVHSNEAGVLMKGNVPTAALLKAVLNKEYGLRIGKVLSHVAVFEIPGYDRLIIVTDAAMNV
SPDLEQKAQIIANAVAVAHAIGIETPKVAPLAAVEIVNPSMPATVDAAALAMMQKRGQIKGCVIDGPLALDNAVSP
KAAKHKRIDSEVAGKADILLVPDIEAGNMLYKSLVYFANAKVGAIIAGAKAPIVLTSRADSAESKLYSLALAVCSA
SK-

Acetate kinase

Coding DNA sequence:

ATGATGATTAAGGGAGTTGTTTGGAATATGGCGAAAATTTTAGCAATTAACGCCGGCAGTTCATCATTAAAGTTTC
AATTATTTGAAATGCCAAGCGAAAAGGTTTTAACAAAAGGAATAGTGGAGCGAATCGGTTTTGATGACGCCATTTT
TACGATCACCGTCAATGGCGAGAAAATTCAAGAAGTCACAGCCATTCCGAACCATGCAGTAGCGGTGAAAATGCTG
CTCGACAAGCTCATGGGGTACAAAATTATTCGCTCATTTGATGAAATCGACGGAATTGGCCACCGCGTTGTGCATG
GTGGCGAAAAGTTCAGTGATTCTGTACTCATTACGGATGAAGTATTAAAACAGATCGAAGAAGTGTCTGAACTTGC
TCCCCTTCATAATCCAGCAAACATTACCGGCATTAAAGCGTTTCAAGAAGTGCTGCCGAATGTGCCTGCCGTGGCG
GTATTTGATACGGCGTTTCATCAAACGATGCCAGAGCAGTCGTTTTTGTATAGTCTCCCATATGAATATTATACAA
AATTCGGCATTCGCAAATACGGCTTCCATGGCACATCACATAAATATGTCACACAACGCGCCGCTGAATTGCTCGG
CCGTCCGATTGAGCAGCTTCGCCTCATTTCTTGCCACCTTGGCAATGGCGCCAGCATCGCTGCAGTAGAAGGGGGA
AAATCGATCGATACGTCCATGGGCTTTACGCCGCTTGCCGGCGTGGCGATGGGAACGCGTTCTGGAAACATTGATC
CGGCGCTTATCCCTTATATTATGGAAAAGACGGGAATGACTGCCGAAGAAGTAGTGGAAGTATTAAATAAGAAAAG
CGGAATGCTTGGCATTTCCGGCATTTCAAGCGACTTGCGCGATTTGGAAAAAGCGGCGGCAGAAGGAAATAAACGT
GCTGAGCTAGCACTTGAAGTATTCGCTAATCGCATTCATAAATATATCGGTTCTTATGCGGCGCGCATGTGCGGTG
TGGATGCGATTATTTTCACTGCTGGAATTGGGGAAAACAGCGAGCTGATCCGTGCCAAAGTATTGCGCGGCCTCGA

ATTTATGGGGGTATACTGGGATCCTTCATTAAATAAAGTACGCGGCAAAGAAGCGTTCATCAGCTATCCGCATTCG
CCGGTAAAAGTGCTTGTCATCCCGACAAATGAAGAAGTCATGATCGCCCGGGATGTTGTCCGGTTGGCGAATATTG
CTTAA

Peptide sequence:

MMIKGVVWNMAKILAINAGSSSLKFQLFEMPSEKVLTKGIVERIGFDDAIFTITVNGEKIQEVTAIPNHAVAVKML
LDKLMGYKIIRSFDEIDGIGHRVVHGGEKFSDSVLITDEVLKQIEEVSELAPLHNPANITGIKAFQEVLPNVPAVA
VFDTAFHQTMPEQSFLYSLPYEYYTKFGIRKYGFHGTSHKYVTQRAAELLGRPIEQLRLISCHLGNGASIAAVEGG
KSIDTSMGFTPLAGVAMGTRSGNIDPALIPYIMEKTGMTAEEVVEVLNKKSGMLGISGISSDLRDLEKAAAEGNKR
AELALEVFANRIHKYIGSYAARMCGVDAIIFTAGIGENSELIRAKVLRGLEFMGVYWDPSLNKVRGKEAFISYPHS
PVKVLVIPTNEEVMIARDVVRLANIA-

Butyrate kinase

Coding DNA sequence:

TTGCAGGAGCAGAAGTTTCGGATTTTAACGATCAATCCTGGATCAACTTCGACAAAAATCGGTGTGTTTGATAATG
AACGCCCCATTTTGGAAAAAACGATCCGCCATGACTTAGAAACGTTGCAGCAATATCCATCGGTTACCGCGCAATA
TGCATTTCGTAAACAGACGATTTTAGATGCGTTGGACGAAGAAGGGCTCGATCTTTCCAAATTAAGCGCTGTCTGC
GGTCGCGGCGGCTTGTTGCGCCCGATTGAAGGAGGAACGTACCGCGTCAATAAGCAGATGATCGAGGATTTGCAAA
AAGGCTATTCCGGCCAGCACGCTTCCAATCTTGGCGGCATTTTAGCCTATGAAATTGCTTCCGCTTTAAATATTCC
CGCGTTTATTGTCGATCCTGTCGTTGTCGATGAACTTGAACCAATTGCCCGCATTTCCGGATTTGCACTCATTGAG
CGAAGAAGCATTTTCCACGCGTTAAACCAAAAGGCGGTGGCGCGCCGCGTCGCCAAACAGCTTGGGCGGCGCTATG
AAGAAGTAAACTTGATCGTTGCCCATATGGGAGGCGGAATCACCGTTGGCGCCCATAAAAAAGGAAGAGTGGTGGA
TGTAAATAACGGGTTGGACGGGGAAGGGCCGTTTGGCCCGGAGCGGGCTGGCACGGTTCCCGCGGGAGATCTTGTT
TCGCTATGTTTCTCAGGAGAATATTACCGCGATGAAATAATGAACATGCTGGTCGGAAAAGGCGGGCTCGTTGGTT
ATCTCGGCACGAATGATGCGGTAAAAGTTGAAAAAATGATCGAAGCGGGAGATGAAAAGGCAAAGCTTGTTTATAG
CGCTATGGCGTATCAAGTAGCAAAAGAAATTGGCGCGGCAAGCGCCGTATTGGCAGGAAATGTCGATGCGATCATT
TTGACAGGCGGGCTTGCATACGGAAAGCAGTTTGTGAAAGAAATTATTGATCGAGTTGACTGGATTGCCGATGTGA
TCGTTCACCCCGGAGAAAATGAACTGCAGGCGCTTGCCGAAGGCGCGCTGCGTGTCTTGCGCGGAGAGGAAAAAGA
GAAAGTATATCCGAGCGGAACATTCGCGACAGCGACCGTTTTGTAA

Peptide sequence:

LQEQKFRILTINPGSTSTKIGVFDNERPILEKTIRHDLETLQQYPSVTAQYAFRKQTILD
ALDEEGLDLSKLSAVCGRGGLLRPIEGGTYRVNKQMIEDLQKGYSGQHASNLGGILAYEI
ASALNIPAFIVDPVVVDELEPIARISGFALIERRSIFHALNQKAVARRVAKQLGRRYEEV
NLIVAHMGGGITVGAHKKGRVVDVNNGLDGEGPFGPERAGTVPAGDLVSLCFSGEYYRDE
IMNMLVGKGGLVGYLGTNDAVKVEKMIEAGDEKAKLVYSAMAYQVAKEIGAASAVLAGNV
DAIILTGGLAYGKQFVKEIIDRVDWIADVIVHPGENELQALAEGALRVLRGEEKEKVYPS
GTFATATVL-

Acetyl-CoA synthetase 1

Coding DNA sequence:

ATGAAATTTTTGTTTTTTAATAAAATATTGAATATTTTTTCTTTTACGATTATTATTGATATTGAGTCTACCTTTA
TCAACAGTGGAGGGGTATATATGCCAAACAAAACGTTCCATTCTTCATTCACCGCTTCTTTTTCTTGGGAGACCGT
GTTTCAACATTACGACTGGAACCCACGTGAACGTTTCAATGTCGCTCATGAAGTATGTGACCGCTATGCAGAAGAT
CCGACTCGCATCGCCCTGTTTTATGAAAACCATTTAGGAGAAAAGAGAACCATCACCTATCGAGAATTGCGGGATT
GGTCGAACCAAATGGCAAATGTGTTCCGCAAACTTGGAGTCAAAAGAGGAGACAGAGTTTGCGCTCTTTTGCCTAA
AAATCCTGCTCTAGTCATTTATATTTTAGCTGCCTGGAAAGTTGGCGCAGTTTATGTTCCGCTATTTACAGCATTT
GGCCCACAGGCGATTGAATACCGGATTAACCATTCTGAGGCAAAAGTGATGTTGACGAATAATGAACAGCGTGCAA
AATTGCCGCCGAGAGAAAAAATACCAACATTAGAGCATATTTTTGTCATTGATGAAACAAATGACCATAACGATTT
ATCTTTTTGGAAAACGCTTTCAGAAGAGTCTGTCGATCATCAGACCGAAGAAACGACTGTGGACGATTTACTGGCT
ATCCAATATACATCCGGATCGACAGGAATGCCGAAAGGAGCCATGTGGTCGCACAATTTGCTCATTAATATATATC
CGTATATGCGATATGCCATCAGTTTGCGGGATGATGATGTATTTTTAGGGGGAGCCGATCCGGGATGGGCTTACGG
GCTGATCTTTTGCACATTCGCGCCAATGTGTTTCGGTGTGCCCATCGTCTTTTACGAAGGACCGTTTAAACCGGAC
ACATACTATTCTCTGATGGAAAAATATCATGTAACGAATTTCACTTACGCTCCGACCGCTTATCGCGCGATGGTCG
CGGCAGGAGCAGAATTGATCAATAAATACAACATCCGCGTGCGGGCAATGAGTTCCGCAGGAGAGCCGCTTAATCC
GGAAGTGATTCGCTTTTTCCAAAAACATTTAGGGGTGACCGTCCACGACCACTACGGTCTGTCGGAAACGTTAATG
CTGATTGGCAATTTCAACGCAATTGACATGGAGATACGCCCTGGTTCGATGGGATGGGTGCTGCCGGGGTTTGAAG
TCGCATTGCTCAATGAAGAAGGCGTTCCGGTAGAAGACGGAAATGTTGGACAAATTGCTTTTAACACGAATTCCAT

TCCAAACGTGTTCAAAGGTTATTGGAAAGATCCTGAAAAAACAGCGGAGAGATTGTCAGGAAGCTGGTTCCTCACC
GGAGATTTGGCAACGAAAGACGCCGACGGATACTTTTGGTTCCAAGGCAGGGCCGATGATATCATCTCCAGTGCCG
GATATCGCATCGGGCCTTTTGAAATTGAGAGTTGCCTCATCGAACATCCAGCGGTAGTGGAAGCAGCGGCCATTGG
AAAGCCAGATCCGTTAAAAGGGGAAATCGTCAAAGCCTTCGTCGTACTGGCAGATGGGTTTACTCCATCGGACGAT
CTGGCGCAAGAATTATCATTATTCGTGAAAAACCGTTTATCGAAGCACGAATACCCTCGCGAAGTGGAATTCGTAG
CAGAACTGCCTAAGACGCCTAGCGGCAAAATTCAGCGATTTATCCTTCGCAATCAAGAAAGCGAAAAACAGAAAAA
CGACTAA

Peptide sequence:

MKFLFFNKILNIFSFTIIIDIESTFINSGGVYMPNKTFHSSFTASFSWETVFQHYDWNPRERFNVAHEVCDRYAED
PTRIALFYENHLGEKRTITYRELRDWSNQMANVFRKLGVKRGDRVCALLPKNPALVIYILAAWKVGAVYVPLFTAF
GPQAIEYRINHSEAKVMLTNNEQRAKLPPREKIPTLEHIFVIDETNDHNDLSFWKTLSEESVDHQTEETTVDDLLA
IQYTSGSTGMPKGAMWSHNLLINIYPYMRYAISLRDDDVFLGGADPGWAYGLIFCTFAPMCFGVPIVFYEGPFKPD
TYYSLMEKYHVTNFTYAPTAYRAMVAAGAELINKYNIRVRAMSSAGEPLNPEVIRFFQKHLGVTVHDHYGLSETLM
LIGNFNAIDMEIRPGSMGWVLPGFEVALLNEEGVPVEDGNVGQIAFNTNSIPNVFKGYWKDPEKTAERLSGSWFLT
GDLATKDADGYFWFQGRADDIISSAGYRIGPFEIESCLIEHPAVVEAAAIGKPDPLKGEIVKAFVVLADGFTPSDD
LAQELSLFVKNRLSKHEYPREVEFVAELPKTPSGKIQRFILRNQESEKQKND-

Acetyl-CoA synthetase 2

Coding DNA sequence:

TTGAACATCATAAAGGAAGATCCAATTTGGTTTCCTTCCGAAAATTATATTGCTGGTTCTCATTTGCAACGATTGA
TTCATCGGCTTGGGTTAAACAACTATGATGAATTATATGAATATAGCGTATTGCATACCGATAAATTTTGGAAAGC
AACGTTGGATGATCTAGGCATTGAATGGTTTGCGCCGTATGAAAAGTTTGTCGATCTATCGGCAGGACCGCAGTGG
CCAAAGTGGTTTGTTGGCGGAAAAATTAACCTTGTTCACAATGCTTTAACGCGTCCTTTGAAACTGGGACGCGGGG
AGCAAACAGCGATAATATGGGAAGACGAATCAGGGATTCGCGAGTCTTTCACTTATAAAGAACTTGATGAACAGAC
ATGCCGTGTCGCAAACGGGTTGCTGGAACTTGGGTTTCGCCGGGGAGACACAATAGGTCTATTCGTTCCGATGATT
CCTGAAGCAGCCATCTTTTTCCTTGCAGCAATTCGTTTGGGGGGGATTGTAGTACCTATTTTTTCTGGATACGGAG
CGGATGCTGTCGCTACTCGTTTAAATGATGCGAAGGCTCGCTTTTTGTTGACAGCAGATGGATTTACAAGGCGGGG
TAAGGTAGTGCCGTTGAAGCGTACAGCTGACGAAGCTGTGGAAAAATGTCCAACTGTCGAACGTGTCATTATGGTT
GAACATCTCAATAATGTTTCCAAAGCAGATCTTATGGAGAAAGATATTATTTGGAAAGATCTTACTGAAGGAAAAG
CAACCACAGCGCCTATAGAAGCCATGGATGCCAACGATCCTTTTATGATTATTTATACATCCGGAACTACTGGCAA
ACCTAAGGGAGCCGTGCATACCCAAGGCGGTTTCCCGCTAAAGACCGTGCAAGATATGGCTCATGTTTTCGATCTA
AGAGCCGGTGAGGTGCTATTTTGGTTGACGGATTTAGGATGGATGGTTGGGCCAGCGGTGATCATTGGTGCGCTAA
CGCTTGGAGCAACCCTAGTATTATACACGGGTGCCCCGGACTACCCTGACCCTGGACGAATTTGGAGTATTGTGGA
AAGAAATAAAGTAACACATTTGGGACTTGCTCCGACCCTCATTCGTTCATTGCGTTCATTTGGTGAAGGACCAATC
CGGAAACGAAATCTCGACACGTTGAGAGTGCTGATTTCCACAGGTGAAGCTTGGGATCTGGAATCTTATCTCTGGT
ATTTTCATGAAGTTGGCAAAGGAAAATTGCCGATTTCAAACTATTCGGGGGGAACGGAAGTATCTGGAGGCATTCT
TGGATGTGTCACTCACCGTCCGATTAAGCCGGCAGGATTTAACACGACTGTTCCAGGAATACAAGCTGCTGTTCTC
GATGACCAAGGACAATCTGTGATTGGCCAAGTAGGCGAATTGGCTGTTTTACGTCCTTTTGTTGGAATGACAAATG
GTTTTTGGAATGACAAAGAACGATATTTGGATTCGTATTGGTCGAAGTTTGAAAATGTGTGGGTCCACGGAGATTG
GGCCGTTCAGGACTCAGAGAATCATTGGTTTCTCTTTGGTCGCTCTGATGATGTGATAAAGGTAGCCGGTAAAAGA
GTTGGACCAGCGGAAATTGAGGCTGCCGCTAGTCAACATCCAGCTGTCTTACAAGCTGCTGCAGTAGGTGTTTCTC
ATCCAGTCAAAGGGGAAGAGGTTGTGCTATTTATAATTCTGCACCCTAATTTCGCTCCAAATGAGGGACTCGCTGA
AGACATTAAAGAGCTGGTAACAAGCGCACTAGGAAAATCATTAAGACCAAGCCGCATTTACTTTGTGAATGATCTG
CCGCGCACGCGTAATGGAAAGGTAATGCGCCGCGTTATCCGGGCCGTGTATCAAGGCCAAAACCCAGGAGATCTTT
CTGCTCTGGAGAATCCAGACATCATTGAGATGATACAAAATTCGACTAAAACACAAAATGGGTAA

Peptide sequence:

MNIIKEDPIWFPSENYIAGSHLQRLIHRLGLNNYDELYEYSVLHTDKFWKATLDDLGIEWFAPYEKFVDLSAGPQW
PKWFVGGKINLVHNALTRPLKLGRGEQTAIIWEDESGIRESFTYKELDEQTCRVANGLLELGFRRGDTIGLFVPMI
PEAAIFFLAAIRLGGIVVPIFSGYGADAVATRLNDAKARFLLTADGFTRRGKVVPLKRTADEAVEKCPTVERVIMV
EHLNNVSKADLMEKDIIWKDLTEGKATTAPIEAMDANDPFMIIYTSGTTGKPKGAVHTQGGFPLKTVQDMAHVFDL
RAGEVLFWLTDLGWMVGPAVIIGALTLGATLVLYTGAPDYPDPGRIWSIVERNKVTHLGLAPTLIRSLRSFGEGPI
RKRNLDTLRVLISTGEAWDLESYLWYFHEVGKGKLPISNYSGGTEVSGGILGCVTHRPIKPAGFNTTVPGIQAAVL
DDQGQSVIGQVGELAVLRPFVGMTNGFWNDKERYLDSYWSKFENVWVHGDWAVQDSENHWFLFGRSDDVIKVAGKR
VGPAEIEAAASQHPAVLQAAAVGVSHPVKGEEVVLFIILHPNFAPNEGLAEDIKELVTSALGKSLRPSRIYFVNDL
PRTRNGKVMRRVIRAVYQGQNPGDLSALENPDIIEMIQNSTKTQNG-

Acetyl-CoA synthetase 3

Coding DNA sequence:

TTGGCGAAAATGAAACGGGAAGATTTAATCGCTCCTGAACGCTATAATTTGACGGTGGAAATCGAGAGGCATGCGC
TGGCTAATCCTGATAAAACCGCATTGAAATGGGAAAGCGAACAAGGGGAAACGCGGGAGATTACATACGGTGATTT
AATCAAACGTGCGAACCGGATTGGAAACGCTTTATTGAAGCAAGGGCTTGAAAAAGGCGATAAAGTGCTTGTCATG
ATTCCTCGCTTAATTGAAGCGTATGAAGTATATTTAGGAGCGTTAAAGGCAGGCTTGGTCGTGATTCCAAGTTCGG
AAATGCTGCGGACGAAAGATTTGCAATACCGCATTTCCCACGGCGAGGCAAAGGCGGTTGTTGCATATGAGCCATA
TGTCGATCAATTTGCGCCAGTGGAAGGCATGGACCATCTCGTGAAATTCATTGTCGGAAAAGAGAAACGGGACGGA
TGGATTCTTTTAGAGGAAGCGATGAAAGCGGAAAGCGATGAACTGGCTACTGCAGACACATCCCGTGATGATATGG
CGTTTTTATCCTATACATCGGGAACGACCGGATATCCGAAAGGAGTCGTTCATTCGCACGGATGGGCATATGCACA
TTTGCGCATTGCCGCGAAGAATTGGTTATGTATTGGAGAAAATGATCTCGTCTGGGCGACCGCCGGCCCAGGCTGG
CAAAAATGGATTTGGAGCCCGTTTTTATCGACGCTTGGTTCTGGTGCAACAGGGTTTGTCTATTATGGACGGTTTG
ATCCGGAAAAATATTTGCAGCTTTTAAGCAAGTATGAAGTAAACGTGCTTTGTTGCACGCCGACGGAGTATCGGCT
GATGGCGAAAGTGCCGAACATCGGTGATTATAATCTCTCGCATCTTCATAGTGCCGTATCCGCCGGGGAGCCGCTT
AACCGTGAAGTGATTGACACGTTTGAGAAATATTTTAATATTCAAGTACGGGACGGATATGGGCAAACGGAAAATA
CGCTGCTTGTCGGCGTCATGAAAGGAATGAAAATCAAACCAGGCTCCATGGGCAAACCAACGCCGGGAAATCGCGT
TGAGATTATTAACGAAAACGGCGAACCGTGCGCCGTCGGGGAAGTTGGCGACATCGCCGTCCATGTGGAAACTCCG
ACGTTATTTAAATATTACTACAAAGATCCAGAACGGACGGCGCTGCAATTCCGAGGCGATTATTATATTACGGGGG
ATAAAGCGAAAAAAGATGAAGATGGATATTTCTGGTTTGAAGGAAGAGGAGATGACATTATCATCAGCTCCGGTTA
TACAATCGGGCCGTTTGAAGTGGAAGATGCGCTTGTCAAGCATCCTTACGTAAAAGAATGCGCTGTCGTCGCCAGT
CCGGATGAAGTGCGCGGCCATGTCGTCAAAGCGTTCATCGTTCTTCGCGAAGGAGTTGATAAAAATGACCCGACGC
TGATTCCAAAGCTGCAGGAGCATGTGAAACAACTTACCGCACCATATAAGTATCCGCGCAAAATTGAATTTGTCGA
CGATTTGCCGAAAACGGCATCGGGAAAAATCCGCCGTGTGGAGCTGCGCGAGCGCGAAATGCGCCTTGCCAAGCGT
CAATAA

Peptide sequence:

MAKMKREDLIAPERYNLTVEIERHALANPDKTALKWESEQGETREITYGDLIKRANRIGNALLKQGLEKGDKVLVM
IPRLIEAYEVYLGALKAGLVVIPSSEMLRTKDLQYRISHGEAKAVVAYEPYVDQFAPVEGMDHLVKFIVGKEKRDG
WILLEEAMKAESDELATADTSRDDMAFLSYTSGTTGYPKGVVHSHGWAYAHLRIAAKNWLCIGENDLVWATAGPGW
QKWIWSPFLSTLGSGATGFVYYGRFDPEKYLQLLSKYEVNVLCCTPTEYRLMAKVPNIGDYNLSHLHSAVSAGEPL
NREVIDTFEKYFNIQVRDGYGQTENTLLVGVMKGMKIKPGSMGKPTPGNRVEIINENGEPCAVGEVGDIAVHVETP
TLFKYYYKDPERTALQFRGDYYITGDKAKKDEDGYFWFEGRGDDIIISSGYTIGPFEVEDALVKHPYVKECAVVAS
PDEVRGHVVKAFIVLREGVDKNDPTLIPKLQEHVKQLTAPYKYPRKIEFVDDLPKTASGKIRRVELREREMRLAKR
Q-

Acetyl-CoA synthetase 4

Coding DNA sequence:

ATGAAAACAGAAGTGCTACCGGTCGTACAAGGGGATTATAACTTAAAAAATTATGAAGAGGTATATAAAAATTTTG
ACTGGAAGGAAGTAGAAAAGGAATTTTCCTGGTTTGAAACAGGACGGGTCAATATGGCATATGAGGCGATTGACCG
CCATGTGGAAACGTTCCGAAAAAATAAAGTGGCGCTTTATTACCGCGATGCGTCGAGAAATGAAAAGTATACGTTT
AAAGAAATGAAAGAATTGTCCAACAAAGTGGCAAACGTATTGAAAGAAGTAGCAGAGATCGAAAAAGGGGACCGTG
TATTTGTTTTTATGCCGCGTTCGCCGGAGCTGTATTTTGCGGTATTAGGAATCATTAAAACAGGCGCGATTGCCGG
GCCGCTGTTTGAGGCGTTTATGGAAGGCGCGGTGCGTGACCGCTTGGAGGATAGCGGGGCGAAGGCGATCGTAACG
ACTCCGGAACTTTTGCCGCGGGTGCCAGTAAACGATCTGCCTGCGCTGAAGTATGTTTTTCTTGTCGGGGACAACA
TTGTTGAAGAAGGTCCGTATATTGATTTAAAGAAACGAATGAATGAAGCAAGCAAGCATTTCGATATTGAATGGGT
CGATCGCCAGGACGGGCTTATTTTGCACTATACATCCGGTTCCACTGGAAAGCCAAAAGGGGTTTTACATGTTCAC
AATGCAATGATTCAGCACTATCAAACGGCAAAATGGGTGCTTGATTTAAAAGAAGATGATATTTATTGGTGCACCG
CTGACCCGGGATGGGTGACAGGAACGTCATATGGCATTTTCGGACCGTGGTTGTGCGGAGCGTCCAACGTCATTAT
TGGGGGACGTTTCAGCCCGGATACGTGGTATCAAGCGATTCAAGATTTTGGCGTGACGGTGTGGTATAGCGCGCCA
ACGGCGTTTCGCATGTTGATGGGCGCCGGAGACGAAATTGTGAAAAAATACGATTTAAGTTCGCTCCGTCATATTC
TTAGCGTTGGAGAACCGTTGAATCCAGAAGTGATTCGCTGGGGAATGAAAGTGTTCCAGCGCCGCATTCATGACAC
ATGGTGGATGACGGAAACAGGTGCTCATCTCATTTGTAACTATCCTTGCATGGAAATCAAGCCAGGCTCGATGGGA
AGGCCGTTTCCTGGCGTTAAAGCAGCCATCATTGATGATCAAGGAAATGAACTGCCGCCATACCGGATGGGGAACT
TGGCGATCAAAAAAGGCTGGCCGTCGATGATGAAAACGATTTGGAACAATCCGCAAAAATATGAATCGTATTTCAT
CGGCGATTGGTACGTTTCCGGCGACTCCGCTTATATGGACGAAGACGGATATTTCTGGTTCCAGGGCCGTGTTGAT
GACGTTATTAATACGTCCGGAGAACGCGTCGGTCCGTTTGAAGTCGAGAGCAAGCTGGTTGAACATCCGGCGGTCG
CGGAAGCAGGCGTCATCGGCAAACCGGATCCGGTGCGCGGGGAAATTATTAAAGCATTTATTTCGCTTCGGGAAGG
ATATGAACCGTCAGAGGAGCTGAAAGAAGACATCCGCCAGTTTGTCAAAAAAGGATTGGCGGCACATGCAGCGCCG
CGGGAAATTGAATTCCGCGACAAACTTCCAAAAACGCGAAGCGGAAAGATCATGCGCCGCGTCCTAAAAGCGTGGG
AACTCAATTTGCCAACGGGGGATTTGTCGACAATGGAAGATTAA

Peptide sequence:

MKTEVLPVVQGDYNLKNYEEVYKNFDWKEVEKEFSWFETGRVNMAYEAIDRHVETFRKNKVALYYRDASRNEKYTF
KEMKELSNKVANVLKEVAEIEKGDRVFVFMPRSPELYFAVLGIIKTGAIAGPLFEAFMEGAVRDRLEDSGAKAIVT
TPELLPRVPVNDLPALKYVFLVGDNIVEEGPYIDLKKRMNEASKHFDIEWVDRQDGLILHYTSGSTGKPKGVLHVH
NAMIQHYQTAKWVLDLKEDDIYWCTADPGWVTGTSYGIFGPWLCGASNVIIGGRFSPDTWYQAIQDFGVTVWYSAP
TAFRMLMGAGDEIVKKYDLSSLRHILSVGEPLNPEVIRWGMKVFQRRIHDTWWMTETGAHLICNYPCMEIKPGSMG
RPFPGVKAAIIDDQGNELPPYRMGNLAIKKGWPSMMKTIWNNPQKYESYFIGDWYVSGDSAYMDEDGYFWFQGRVD
DVINTSGERVGPFEVESKLVEHPAVAEAGVIGKPDPVRGEIIKAFISLREGYEPSEELKEDIRQFVKKGLAAHAAP
REIEFRDKLPKTRSGKIMRRVLKAWELNLPTGDLSTMED-

Lysine acetylase, AcuA

Coding DNA sequence:

ATGAAAACAGAAGAAGAAATTCCAGGTGGTGACCTAATGGAACATAAAAAGACATATAATGCAAAAGAATTAAAGA
CCCCGAAGGGAACGCTGATTATTGAAGGCCCGGTACCCGCAGAAAAACTAGCAAGCTACGAATTTCACCACGATTT
GACATCATTCCGCCAGCCACCGCAGCAGCATCAAGCGCTCATTGAGATCGCTAAACTCCCGGAAGGAAGAATCATC
ATCGCCCGCCATCATCAAACAATCGTCGGTTATGTTACATTTCTTTACCCTGATCCGCTGGAACGATGGTCGGAAG
GAAATATGGAAAATTTAATTGAATTGGGTGCGATTGAAGTCATTCCGGAGTTTCGCGGCTATGGCGTGGGAAAAAA
CTTGCTGATTGTATCCATGATGGATGATGCGATGGAAGATTACATTATTATTACGACAGAATATTACTGGCATTGG
GATTTAAAAGGAACAGGATTAAACGTATGGGAGTACAGAAAAGTAATGGAAAAAATGATGAACGCTGGCGGGCTTG
TTTGGTATGCAACAGATGATCCCGAAATTTGCTCTCATCCAGCTAACTGCCTGATGGTCCGCATCGGGAAACGGGT
CGATCAAGAATCCATCCAAAAATTTGACCGCCTTCGCTTCATGAACCGGCACATGTATTAA

Peptide sequence:

MKTEEEIPGGDLMEHKKTYNAKELKTPKGTLIIEGPVPAEKLASYEFHHDLTSFRQPPQQHQALIEIAKLPEGRII
IARHHQTIVGYVTFLYPDPLERWSEGNMENLIELGAIEVIPEFRGYGVGKNLLIVSMMDDAMEDYIIITTEYYWHW
DLKGTGLNVWEYRKVMEKMMNAGGLVWYATDDPEICSHPANCLMVRIGKRVDQESIQKFDRLRFMNRHMY-

Lysine deacetylase, AcuC

Coding DNA sequence:

ATGGCCGAATTTGCCGGGGGTTTCGTCATGAAAAAAGATTGCGTATTTATTTACAGCGAAGATTTCCTCCGATATA
AATTTCATGACGCGCACCCGTTTAACCAGCTGCGTGTCAAGCTGACATACGACTTATTGCGCGCAATGAACGCACT
GGAAGATGAGCAGATCATCGAACCAAGAATCGCGACGGACGAGGAACTCGCCTTGATTCATGATCAAACATATATT
GAAGCGGTCAAAGCGGCCGGAAAAGGGCAGCTGCCTGAGCATACCGCCTTAAACTACGGGCTGGGAACGGAAGATA
CGCCAATTTTCCCAAATATGCATGAAGCAAGCGCATTGTTGGTGGGAAGCACATTAACCGCTGTCGATTATGTGCT
ATCCGGAAAAGCAAAGCATGCGCTAAGTTTGGGCGGCGGGCTTCACCACGGCTTCCGCGGCAAAGCATCCGGATTT
TGCGTTTACAATGATAGCGCTGTCGCCATTAAATACATTCAAGAAAAATACGGGCTTCGCGTTCTTTATGTTGACA
CGGATGCCCATCACGGAGACGGCGTCCAATGGGCATTTTATGATGATCCTAACGTATGCACTTTTTCCATTCACGA
AACAGGCCGCTATCTGTTCCCGGGAACAGGAAACATTAACGAGCGCGGGCACGGAGCGGGATACGGCTATTCCTTT
AACATTCCTGTCGATGCGTTTACAGAAGATGAATCCTGGCTTGACGCTTATACGCAAGCACTTCGGGAAATCGCTG
ACTTTTTCCGCCCTGATGTGATCTTGACGCAAAACGGGGCGGACGCCCACTATTACGACCCGCTTACCCATTTGTC
GTTAACAATGAAAACATACCGGGAAATTCCAAAGCTTGCACATCAAATTGCCCACGAATATTGCGACGGGCGTTGG
ATCGCGGTCGGCGGCGGCGGTTATGATATTTGGCGCGTCGTCCCACGGGCATGGGCGTTTATTTGGTTGGAAATGA
CAGAACAATCAAATGTTTCTGGCAAATTACCAGAATCATGGCTCAAAAAATGGCAATCTCTCTCCCCTATCCCGCT
CCCGCAAGAATGGGACGATCCGGAATCGCTATATCCGCCGATTCCGCGCAAAGCGGAAATTACCGAAAAAAATGCG
CAAACCGTCGAAAAAGCGTTATATTTAATTCGCAGCCAACGGCAACAAACAAACGGATACCAGTCCAACCATTTCT
AA

Peptide sequence:

MAEFAGGFVMKKDCVFIYSEDFLRYKFHDAHPFNQLRVKLTYDLLRAMNALEDEQIIEPRIATDEELALIHDQTYI
EAVKAAGKGQLPEHTALNYGLGTEDTPIFPNMHEASALLVGSTLTAVDYVLSGKAKHALSLGGGLHHGFRGKASGF
CVYNDSAVAIKYIQEKYGLRVLYVDTDAHHGDGVQWAFYDDPNVCTFSIHETGRYLFPGTGNINERGHGAGYGYSF
NIPVDAFTEDESWLDAYTQALREIADFFRPDVILTQNGADAHYYDPLTHLSLTMKTYREIPKLAHQIAHEYCDGRW
IAVGGGGYDIWRVVPRAWAFIWLEMTEQSNVSGKLPESWLKKWQSLSPIPLPQEWDDPESLYPPIPRKAEITEKNA
QTVEKALYLIRSQRQQTNGYQSNHF-

Frag#	Res#	Sequence	Theor (Bo)	[M+H]
▪T1	1-13	(-)MAMLNLVAFLYNK(S)	1526.80	1527.81
T2	14-19	(K)SILVYR(G)	749.44	750.45
T3	20-25	(R)GISNIR(E)	658.38	659.38
T4	26-39	(R)EDSFVSSDLFSTLK(E)	1573.75	1574.76
T5	40-41	(K)EK(I)	275.15	276.16
T6	42-45	(K)IAGK(Q)	387.25	388.26
T7	46-47	(K)QR(K)	302.17	303.18
T8	48-48	(R)K(I)	146.11	147.11
T9	49-58	(K)IVFPEGLDER(I)	1173.60	1174.61
T10	59-65	(R)ILTAVSR(L)	758.47	759.47
T11	66-84	(R)LANEQIVTPIVIGNEEAVK(Q)	2036.12	2037.12
T12	85-86	(K)QK(A)	274.16	275.17
T13	87-110	(K)ASELGLTLPNVEIIDPHQYGEMDK(L)	2668.31	2669.31
T14	111-118	(K)LVAAFVER(R)	903.52	904.53
T15	119-119	(R)R(K)	174.11	175.12
T16	120-120	(R)K(G)	146.11	147.11
T17	121-122	(K)GK(V)	203.13	204.13
T18	123-129	(K)VTEEAAR(K)	774.39	775.40
T19	130-130	(R)K(L)	146.11	147.11
T20	131-147	(K)LLLDENYFGTMLVYMDK(A)	2063.99	2065.00
T21	148-171	(K)AHGLVSGAAHSTADTVRPALQIIK(T)	2412.32	2413.33
T22	172-173	(K)TK(Q)	247.15	248.16
T23	174-177	(K)QGVR(K)	458.26	459.27
T24	178-178	(R)K(T)	146.11	147.11
T25	179-187	(K)TSGVFIMVR(G)	1008.54	1009.55
T26	188-191	(R)GDEK(Y)	447.20	448.20
T27	192-220	(K)YVFADCAINIAPDSQDLAEIAVESANTAK(M)	3038.45	3039.46
T28	221-227	(K)MFDIEPR(V)	906.43	907.43
T29	228-236	(R)VAMLSFSTK(G)	982.52	983.52
T30	237-240	(K)GSAK(S)	361.20	362.20
T31	241-246	(K)SPETEK(V)	689.32	690.33
T32	247-252	(K)VVEAVR(L)	671.40	672.40
T33	253-255	(R)LAK(E)	330.23	331.23
T34	256-279	(K)EMAPDLVLDGEFQFDAAFVPSVAK(K)	2595.26	2596.26
T35	280-280	(K)K(K)	146.11	147.11
T36	281-281	(K)K(A)	146.11	147.11
T37	282-307	(K)APDSVIQGDAMVFIFPSLEAGNIGYK(I)	2721.36	2722.37
T38	308-311	(K)IAQR(L)	486.29	487.30
T39	312-334	(R)LGNFEAVGPILQGLNKPVNDLSR(G)	2450.33	2451.34
T40	335-343	(R)GCNAEDVYK(L)	997.42	998.43
T41	344-353	(K)LTLITAAQSL(-)	1029.61	1030.61

Unique peptides for full-length PTA

Figure 10.1 Predicted trypsin digest of full-length PTA. Peptide fragments unique to full-length protein indicated.

Frag#	Res#	Sequence	Theor (Bo)	[M+H]
T1	1-10	(-)VSSDLFSTLK(E)	1095.58	1096.59
T2	11-12	(K)EK(I)	275.15	276.16
T3	13-16	(K)IAGK(Q)	387.25	388.26
T4	17-18	(K)QR(K)	302.17	303.18
T5	19-19	(R)K(I)	146.11	147.11
T6	20-29	(K)IVFPEGLDER(I)	1173.60	1174.61
T7	30-36	(R)ILTAVSR(L)	758.47	759.47
T8	37-55	(R)LANEQIVTPIVIGNEEA VK(Q)	2036.12	2037.12
T9	56-57	(K)QK(A)	274.16	275.17
T10	58-81	(K)ASELGLTLPNVEIIDPH QYGEMDK(L)	2668.31	2669.31
T11	82-89	(K)LVAAFVER(R)	903.52	904.53
T12	90-90	(R)R(K)	174.11	175.12
T13	91-91	(R)K(G)	146.11	147.11
T14	92-93	(K)GK(V)	203.13	204.13
T15	94-100	(K)VTEEAAR(K)	774.39	775.40
T16	101-101	(R)K(L)	146.11	147.11
T17	102-118	(K)LLLDENYFGTMLVYMDK (A)	2063.99	2065.00
T18	119-142	(K)AHGLVSGAAHSTADTVR PALQIIK(T)	2412.32	2413.33
T19	143-144	(K)TK(Q)	247.15	248.16
T20	145-148	(K)QGVR(K)	458.26	459.27
T21	149-149	(R)K(T)	146.11	147.11
T22	150-158	(K)TSGVFIMVR(G)	1008.54	1009.55
T23	159-162	(R)GDEK(Y)	447.20	448.20
T24	163-191	(K)YVFADCAINIAPDSQDL AEIAVESANTAK(M)	3038.45	3039.46
T25	192-198	(K)MFDIEPR(V)	906.43	907.43
T26	199-207	(R)VAMLSFSTK(G)	982.52	983.52
T27	208-211	(K)GSAK(S)	361.20	362.20
T28	212-217	(K)SPETEK(V)	689.32	690.33
T29	218-223	(K)VVEAVR(L)	671.40	672.40
T30	224-226	(R)LAK(E)	330.23	331.23
T31	227-250	(K)EMAPDLVLDGEFQFDAA FVPSVAK(K)	2595.26	2596.26
T32	251-251	(K)K(K)	146.11	147.11
T33	252-252	(K)K(A)	146.11	147.11
T34	253-278	(K)APDSVIQGDAIVFIFPS LEAGNIGYK(I)	2721.36	2722.37
T35	279-282	(K)IAQR(L)	486.29	487.30
T36	283-305	(R)LGNFEAVGPILQGLNKP VNDLSR(G)	2450.33	2451.34
T37	306-314	(R)GCNAEDVYK(L)	997.42	998.43
T38	315-324	(K)LTLITAAQSL(-)	1029.61	1030.61

Unique peptides for truncated PTA

Figure 10.2 Predicted trypsin digest of truncated PTA. Peptide fragments unique to truncated protein indicated.

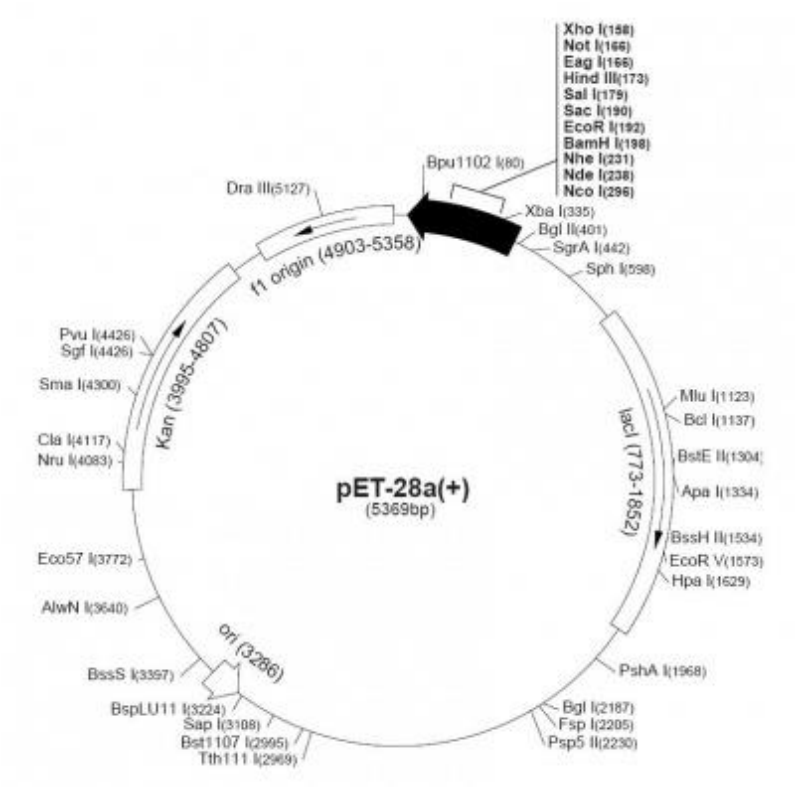


Figure 10.3 Vector map of pET-28a(+).

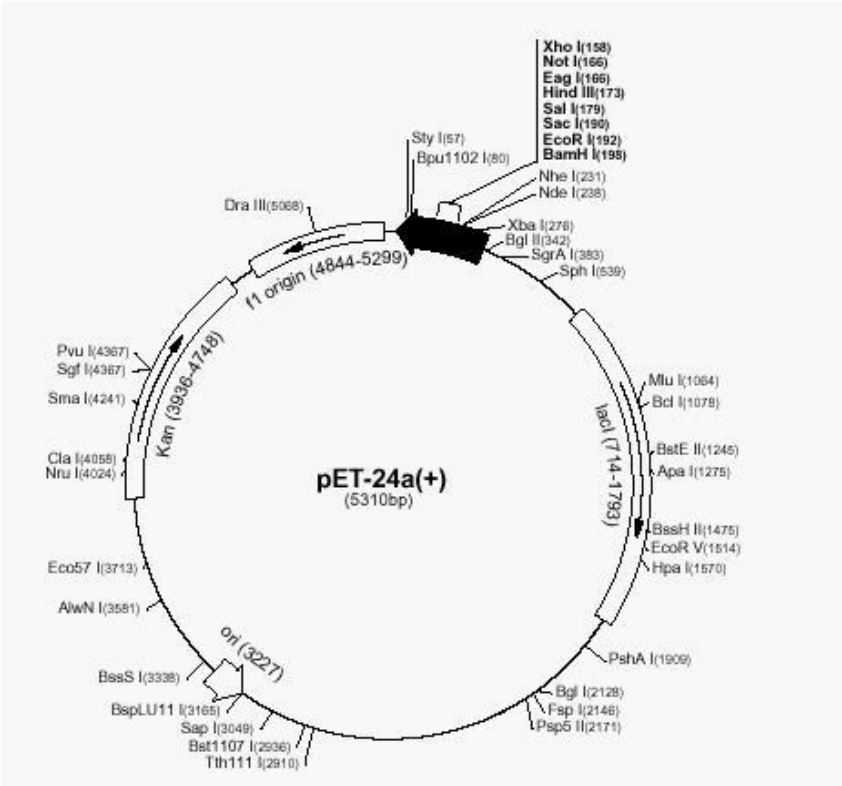


Figure 10.4 Vector map of pET-24a(+).

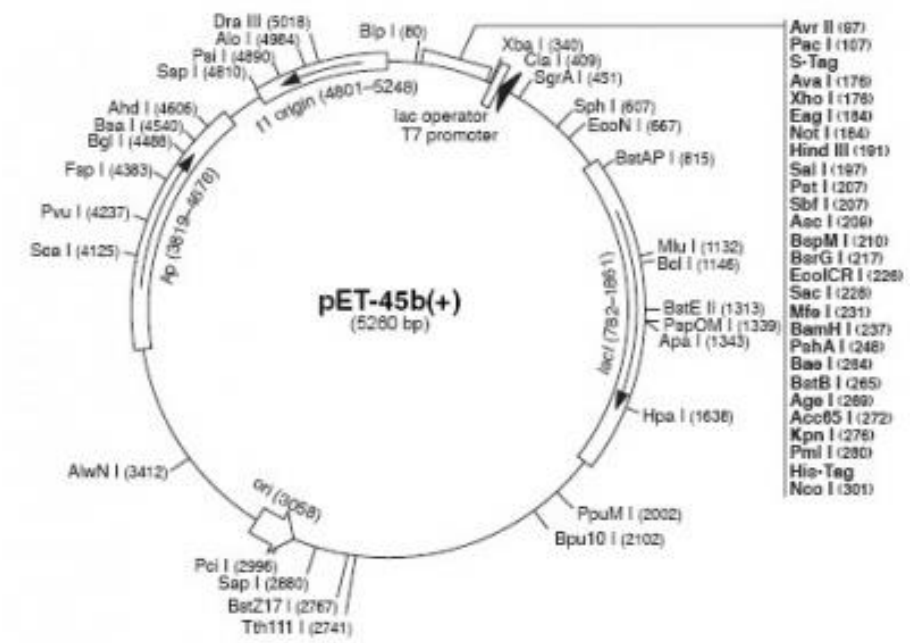


Figure 10.5 Vector map of pET-45b(+).

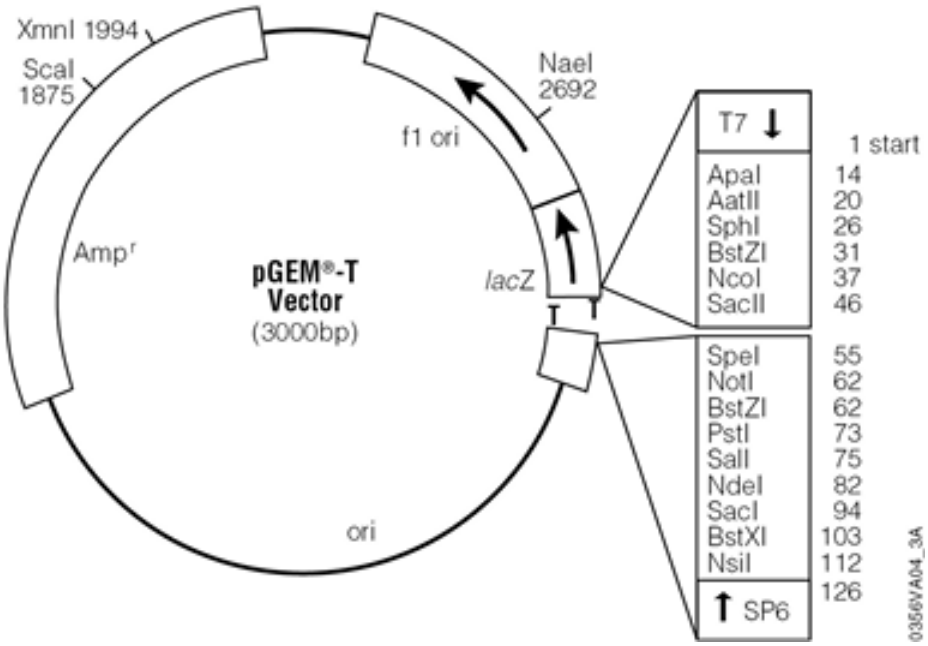


Figure 10.6 Vector map of pGEM®-T.